Unravelling the sex-specificity in Alzheimer's disease risk and its phenotypic variability through brain imaging, Apolipoprotein E polymorphism, and parent-of-origin effects

Chloé Savignac
Integrated Program in Neuroscience
McGill University, Montreal

December 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

First published on December 2024

Table of Contents

Abstract:

Amidst the escalating global challenge posed by Alzheimer's disease (AD), this comprehensive exploration delves into the intricate interplay of genetic, environmental, and sex-specific factors in shaping AD susceptibility and heterogeneity. We examined the manifestations of AD family risk in two extensive epidemiological cohorts, with a specific focus on the impact of the Apolipoprotein E (APOE) isoforms, notably the protective ε2 and deleterious ε4 variants, in shaping AD susceptibility in at-risk males and females. By employing a tailored analytical framework, our objective was to untangle sex biases within AD-related phenotypes as they manifest in distinct subregions of the hippocampus (HC) and default network (DN). Leveraging the robust UK Biobank imaging cohort, we performed a rigorous comparative analysis of brain imaging outcomes associated with ε2 and ε4, revealing discernible effects on brain structure and phenotypic traits. Our population-based approach unveils sex biases in the interaction between ε2 and HC-DN co-variation, impacting both fixed (e.g., AD family history) and modifiable (e.g., social engagement, physical activity) risk indicators. No similar interaction patterns were observed with the commonly studied *APOE* ε4. A complementary investigation into AD genealogy within the Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD (PREVENT-AD) cohort provided further support for the presence of sex bias among individuals carrying the ε2 allele, especially concerning cardiovascular and cognitive risk indicators. Neuroanatomical patterns in HC and DN subregions, influenced by ε2 vs. ε4 polymorphism, highlighted substantial structural variation linked to maternal vs. paternal AD lineage in subregions from which the fornix white-matter tract originates. Across these two prospective cohorts, the ε2 allele stood out as a considerable driver modulating sex differences in AD risk indicators and their neuroanatomical underpinnings. Our cross-generational approach accentuates the need to explore and optimize the relatively less-studied protective mechanisms mediated by *APOE* ε2.

Au cœur du défi mondial complexe posé par la maladie d'Alzheimer (MA), nous entamons une exploration compréhensive de l'interaction entre les facteurs génétiques, environnementaux et spécifiques au sexe, sous-jacents à l'hétérogénéité de la MA. Pour ce faire, nous avons examiné les manifestations du risque familial pour la MA dans deux vastes cohortes épidémiologiques, en nous concentrant spécifiquement sur l'impact des isoformes de l'Apolipoprotein E (APOE),

notamment les variants protecteur ε2 et délétère ε4, dans la formation de la sensibilité à la MA chez les hommes et les femmes à risque. En utilisant un cadre analytique conçu sur mesure, notre objectif était de démêler les biais sexuels dans les phénotypes liés à la MA tels qu'ils se manifestent dans des sous-régions distinctes de l'hippocampe (HC) et du réseau du mode par défaut (MPD). En tirant parti de la riche cohorte d'imagerie cérébrale de la UK Biobank, nous avons effectué une analyse comparative rigoureuse des variations structurelles associées aux allèles ε2 et ε4, révélant des effets perceptibles sur les traits phénotypiques et la structure du cerveau. Notre approche à l'échelle de la population a révélé des biais sexuels dans l'interaction entre l'allèle ε2 et la covariation HC-MPD, ayant un impact à la fois sur des facteurs de risque fixes (c.-à-d. des antécédents familiaux de MA) et modifiables (par exemple, l'engagement social et l'activité physique). Aucune interaction n'a été observée avec l'allèle ε4 couramment étudié. Une enquête complémentaire sur la généalogie de la MA au sein de la cohorte Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD (PREVENT-AD) a renforcé la présence d'un biais sexuel chez les individus porteurs de l'allèle ε2, en particulier concernant les indicateurs de risque cardiovasculaire et cognitif. Des altérations neuroanatomiques dans les sous-régions de l'HC et du MPD, influencées par le polymorphisme de ε2 vs ε4, ont mis en évidence une variation structurelle substantielle liée au risque maternel et paternel de MA dans les sous-régions d'où proviennent les fibres du fornix. Dans ces deux cohortes prospectives, l'allèle ε2 s'est imposé comme un moteur considérable modulant les différences entre les hommes et les femmes dans les indicateurs de risque de MA ainsi que leurs fondements neuroanatomiques. Notre approche intergénérationnelle accentue la nécessité d'explorer et d'optimiser les mécanismes de protection relativement moins étudiés médiés par APOE ε2.

Acknowledgements:

I would like to begin by thanking Dr. Danilo Bzdok for supervising this work and continuously providing opportunities for me to improve as a scientist in both technical and non-technical aspects. I am also grateful for committee members Dr. Mallar Chakravarty and Dr. Judes Poirier, who provided advice and guidance throughout data analysis, results interpretation, and preparation of the thesis. I also wish to thank Dr. Judes Poirier, Dr. Sylvia Villeneuve, Dr. AmanPreet Badhwar, Dr. Maiya Geddes and Mr. Frédéric St-Onge for sharing their knowledge of Alzheimer's disease and advising on the clinical interpretation of our results. Thank you to Dr. Sarah A. Gagliano Taliun and Dr. Sali Farhan for reporting on genetic markers and their significance in the context of Alzheimer's disease. I also wish to thank Dr. Yasser Iturria Medina, Dr. R. Nathan Spreng, and Mr. Chris Zajner for assisting in writing the manuscripts. A special thank you to Ms. Kimia Shafighi, Ms. Karin Saltoun, and Mr. Vaibhav Sharma for providing support in data analysis and preprocessing, as well as for being amazing colleagues. I am grateful to the whole Bzdok lab for their support throughout the completion of my master's degree, and I could not have wished for a better team. I would also like to thank the PREVENT-AD Research Group, the Centre for Studies on Prevention of Alzheimer's Disease (StoP-AD), its director Dr. Judes Poirier, and associate director Dr. Sylvia Villeneuve for sharing their data and guiding us through the PREVENT-AD data resource, facilitating the completion of the second manuscript. The PREVENT-AD data resource was made possible through the generous support of the Canadian Institute on Health Research (CIHR), the Fonds de la Recherche du Québec en Santé (FRQS) and the J.L. Levesque Foundation. I am sincerely grateful for their contributions to our work. I would also like to thank the Brain Canada Foundation, the National Institutes of Health (NIH), the Canadian Institute on Health Research (CIHR), the Healthy Brains Healthy Lives initiative, Google, and the CIFAR Artificial Intelligence Chairs program for supporting Dr. Danilo Bzdok and the Bzdok lab, providing students with the means and resources to conduct state-of-the-art data analyses with cutting-edge equipment. Lastly, I am incredibly grateful to the Canadian Institutes of Health Research (CIHR; Alexander Graham Bell Graduate Scholarship-Master's) and the Fonds de recherche du Québec en Nature et technologies (FRQNT; Bourses de formation à la maîtrise) for awarding me funding throughout my master's degree and believing in my research potential.

Contribution of Authors

Chapter 1: Chapter 1: APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementias

Chloé Savignac and Danilo Bzdok conceptualized the study. Chloé Savignac analyzed the data and performed the statistical analyses. Chloé Savignac and Chris Zajner interpreted the results. Chloé Savignac, Karin Saltoun, Kimia Shafighi, and Vaibhav Sharma worked on data preprocessing and the design of analytic pipelines. Danilo Bzdok supervised the project. Chloé Savignac, Sylvia Villeneuve, AmanPreet Badhwar, Karin Saltoun, Kimia Shafighi, Chris Zajner, Vaibhav Sharma, Sarah A Gagliano Taliun, Sali Farhan, Judes Poirier, and Danilo Bzdok contributed to the writing of the manuscript.

Chapter 2: Parent-of-origin effects in Alzheimer's liability dissociate neurocognitive and cardiovascular traits in at-risk individuals

Chloé Savignac and Danilo Bzdok conceptualized the study. Chloé Savignac analyzed the data and performed the statistical analyses. Chloé Savignac worked on data preprocessing and the design of analytic pipelines. Danilo Bzdok supervised the project. Chloé Savignac, Frédéric St-Onge, Sylvia Villeneuve, AmanPreet Badhwar, Sarah A. Gagliano Taliun, Sali Farhan, Maiya Geddes, Yasser Iturria Medina, Judes Poirier, R. Nathan Spreng, and Danilo Bzdok contributed to the writing of the manuscript and interpretation of the results.

List of Abbreviations

AD	Alzheimer's disease
ADRD	Alzheimer's disease and related dementias
AG	Angular gyrus
APOE	Apolipoprotein E
Aβ	Amyloid-β
BchE	Butyrylcholinesterase
BDNF	Brain-derived neurotrophic factor
BIC	McConnell Brain Imaging Centre
BP	Bloop pressure
CA	Cornu Ammonis
CAIDE	Cardiovascular risk factors, aging, and incidence of dementia
CCA	Canonical correlation analysis
CDK5RAP2	CDK5 regulatory subunit associated protein 2
CIFAR	Canadian Institute for Advanced Research
CIHR	Canadian Institutes of Health Research
CNS	Central nervous system
$cPLA_2$	cytosolic phospholipase A2
CRIUGM	Centre de recherche de l'Institut universitaire de gériatrie de Montréal
CSF	Cerebrospinal fluid
DG	Dentate gyrus
dmPFC	Dorsomedial prefrontal cortex
DN	Default network
FAST	FMRIB's Automated Segmentation Tool
FDR	False discovery rate
FRQNT	Fonds de recherche du Québec en Nature et technologies
FRQS	Fonds de la Recherche du Québec en Santé
FUNPACK	FMRIB UKB Normalisation, Parsing And Cleaning Kit
GC-ML-DG	Granule cell layer and molecular layer of the dentate gyrus
GDC	Gradient distortion correction

HATA	Hippocampus-amygdala transition area
HbA1c	Hemoglobin A1C
HC	Hippocampus
HDL	High-density lipoprotein
HMGCR	3-hydroxy-3-methylglutaryl-coenzyme A reductase
HPA	Hypothalamus-pituitary-adrenal
IGF-1	Insulin-like growth factor 1
IL-15	Interleukin-15
IMT	Intima-medial thickness
IP	Intermediate phenotype
IPL	Inferior parietal lobule
LDL	Low-density lipoprotein
MA	Maladie d'Alzheimer
MCI	Mild cognitive impairment
Mila	Quebec Artificial Intelligence Institute
ML	Molecular layer
MNI	Montreal neurological institute
MoCA	Montreal cognitive assessment
MPD	Réseau du mode par défaut
MRI	Magnetic resonance imaging
mtDNA	Mitochondrial DNA
MTG	Middle temporal gyrus
MTL	Middle temporal lobe
MTS	Middle temporal sulcus
NHS	National Health Service
NIH	National Institutes of Health
NO_2	Nitrogen dioxide
Par	Parietal cortex
Para	Parasubiculum
PCA	Principle component analysis
PCC	Posterior cingulate cortex

PCu	Precuneus
pCUNPCC	Precuneus/posterior cingulate cortex
PFC	Prefrontal cortex
PHESANT	PHEnome Scan ANalysis Tool
PLS	Partial least square
PLS-R	PLS-regression
PM_{10}	Particulate matter of 10 μm or less in diameter
$PM_{2.5}$	Particulate matter of 2.5 μm or less in diameter
PPP2r1A	Protein phosphatase 2 scaffold subunit Alpha
Pre-SMA	Pre-supplementary motor area
PREVENT-AD	Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD
PrS	Presubiculum
RBANS	Repeatable battery for assessment of neuropsychological status
RSC	Retrosplenial cortex
SMG	Supramarginal gyrus
SNP	Single nucleotide polymorphism
SPL	Superior parietal lobule
STG	Superior temporal gyrus
StoP-AD	Centre for Studies on Prevention of Alzheimer's Disease
STS	Superior temporal sulcus
SVD	Singular value decomposition
Temp	Temporal cortex
TLR4	Toll-like receptor 4
TPJ	Temporo-parietal junction
TSH	Thyroid stimulating hormone
UKB	UK Biobank
VEGF	Vascular endothelial growth factor
vlPFC	Ventrolateral prefrontal cortex
vmPFC	Ventromedial prefrontal cortex
vPFC/PFCv	Ventral prefrontal cortex

Introduction

The rising number of dementia cases worldwide is a major public health crisis that affects 50 million people around the globe and generates annual costs of >1 trillion USD$ [1]. The number of individuals living with Alzheimer's disease and related dementia (ADRD) is estimated to reach 250 million by 2050 [1]. This three-fold increase is partly attributable to higher longevity and improved quality of life in middle-income countries [2]. While age remains a critical risk factor for both genders[1], women appear to exhibit a higher susceptibility to the disease beyond the risk attributable to their longer life expectancy [3]. Given that women constitute the majority of dementia caregivers and patients globally, the escalating prevalence of ADRD cases is poised to exert more significant consequences on women compared to men [4]. The reasons for gender and sex disparities in dementia care and incidence are complex and multifaceted, involving both societal and biological factors. At the societal level, dementia diagnosis in women is often delayed as symptoms tend to be overlooked or misattributed to other age-related changes [5, 6]. Consequently, at the time of dementia diagnosis, the severity of symptoms tends to be higher in women than men, potentially influencing the perceived trajectory and symptomology of ADRD across genders [7]. Biological variability between males and females may stem from hormonal and neuro-inflammatory changes in the aging brain, possibly interacting with underlying genetic predispositions [8]. Recent advancements in population-based cohort designs, coupled with the ability to track disease progression in healthy older adults prospectively, have opened new avenues to investigate sex biases in ADRD risk across hundreds to thousands of biological markers. These large-scale prospective studies allow for the reliable assessment of the phenotypic variability in ADRD risk attributable to sex, providing valuable insights into potential sex biases in disease onset, progression, and heritability.

Exploring the biological underpinnings of sex biases in Alzheimer's disease (AD) pathogenesis, researchers have probed the influence of sex hormones, particularly estrogen, on neurovascular functions [9]. Estrogen is thought to exert neuroprotective effects by promoting

[1] We will use the term 'gender' to discuss societal distinctions between women and men, and we will use the term 'sex' to specifically refer to the biological, hormonal, and genetic differences observed in the aging male and female brains.

synaptic plasticity and neuronal survival in brain regions with a high concentration of steroid receptors such as in the hippocampus (HC) [10, 11]. The HC, particularly the pyramidal cells of the CA1 subfield, is known to show early alteration along the AD continuum at least since the late 1990s [12]. The CA1 and subiculum subregions serve as the primary origin of most efferent pathways from the HC to the broader cortical regions [13]. The dentate gyrus (DG) and CA4 subregions primarily project within the HC to the pyramidal cells of CA3, which, in turn, project to the CA1 subregion [14]. Evidence from rodent models suggests that the regulation of adult neurogenesis in the DG and CA4 subregions is influenced by adrenal and gonadal steroids, with a particular emphasis on the role of estrogens. [15-18]. Estrogen promotes and sustains the utilization of glucose as the primary fuel source of the brain by increasing key enzymes in the glycolytic pathway which in turn limits brain fatty acid ketosis [19]. Following menopause, estrogen decline could lead to elevated ADRD susceptibility, by potentiating age-related changes in neuroimmune and metabolic functions [8]. The fluctuation of estrogen levels in post-menopausal adult females has been shown to influence inflammation and cholesterol markers [20]. Furthermore, brain regions that exhibit a notable decline in cholesterol synthesis with aging, such as the HC, display a significant concentration of steroid receptors [10, 21]. The changes in membrane lipid composition observed in the HC of individuals with AD, which are linked to a decrease in cholesterol levels, underscore the importance of lipid homeostasis in the pathophysiology of AD [22, 23]. The reduction in estrogen levels among postmenopausal females might thus contribute to age-related alteration in cholesterol profiles, thereby exacerbating sex biases in AD pathogenesis.

The apolipoprotein E (APOE)[2] protein is central to the intricate web of connections between cholesterol metabolism and AD pathology. APOE plays a crucial role in lipid metabolism and cholesterol synthesis in the central nervous system (CNS) [24, 25]. The ε2 allele is the phylogenetically youngest APOE variant and is thought to have emerged around 80,000 years ago from an arginine-to-cysteine substitution on ε3 [26]. The transition from the ancestral ε4 to the now more prevalent ε3 haplotype can be traced back to approximately 200,000 years ago—though this estimate remains subject to debate [26]. Despite a single amino acid exchange differentiating

[2] In accordance with conventional gene nomenclature, it is customary to italicize gene names while leaving protein names non-italicized.

ε2 from ε3, and ε3 from ε4, these subsequent mutations give rise to significant alterations in the blood lipid profile. High-density lipoproteins (HDL) are potentially favoured over low-density lipoproteins (LDL) with the ε2 and ε3 variants [27]. Among all haplotypes, the ε2 variant exhibits the lowest affinity for LDL receptors [28]. This specific trait potentially contributes to a more efficient clearance of amyloid-β (Aβ) deposition, a hallmark of AD pathophysiology, from the brain of ε2 carriers [29-32]. In contrast, the *APOE* ε4 allele has demonstrated a strengthened binding avidity to the Aβ peptide [33]. This characteristic potentially contributes to a heightened deposition of amyloid plaques in the brain of ε4 carriers [30-32]. As is the case for the regulation of cholesterol balance, the impact of the ε4 allele on AD risk and pathophysiology is not without sex-specific intricacies.

At the age of 65, females have an almost twofold higher remaining lifetime risk of developing AD compared to males [34]. The incidence rate is further elevated for female carriers of the ε4 allele [35]. The association between lipid metabolism and estrogen may underlie the increased risk of AD in female carriers of the *APOE* ε4 allele. In the aging female brain, a decline in mitochondrial respiration and an increase in H_2O_2 production are thought to promote the shift to ketone metabolism [36, 37]. H_2O_2 production activates an astrocyte-mediated ketogenic pathway through cytosolic phospholipase A2 ($cPLA_2$) [37]. Astrocyte reactivity in white matter tracts of the HC in reproductively aging female mice is thought to be greatest in the fimbria and to co-occur with $cPLA_2$ labelling [37]. The fimbria is where the axon bundles branching from the pyramidal cells of the CA1 and subiculum subfields converge to form the fornix, which is the major efferent path from the HC to the brain's default network (DN) [13, 38]. The DN is for most parts composed of phylogenetically younger brain regions which together consume some of the highest oxygen levels in the entire brain [39]. Evidence from healthy individuals and ε4 carriers suggests that regions of late myelination are particularly susceptible to age-related degradation [40, 41]. In humans carrying the *APOE* ε4 allele, a decrease in glucose metabolism was observed in regions of the DN, such as the precuneus (PCu), the posterior cingulate cortex (PCC), the temporoparietal junction (TPJ), and the dorsolateral prefrontal cortex (dlPFc) [42, 43]. Critical hubs within the DN that exhibit diminished glucose metabolism as part of the aging process, notably the PCu and PCC, have been identified as early sites of Aβ accumulation [44]. The PCC and PCu are also thought to be particularly influenced by the interaction of *APOE* and sex, such

that females carrying the ε4 allele exhibited a pronounced reduction in functional connectivity in these exact subregions of the DN [45]. A plausible connection between DN hypoactivity among older female carriers of the ε4 allele and changes in brain glucose metabolism may be attributed to an increased reliance on ketosis following the decline in estrogen observed during menopause. This altered metabolic burden has the potential to disrupt the brain's lipid profile, presenting a contributing factor to the development of age-related neurodegenerative disorders, including AD. The interplay between the vulnerability associated with late myelination in regions like the DN, which is typical in aging, and the distinct metabolic feature of the female brain holds promise for understanding the mechanisms that contribute to sex biases in the risk and pathogenesis of ADRD.

Despite variations in amino acid sequences and protein sizes, the APOE protein is present in terrestrial and marine vertebrates such as mammals, reptiles, and fish [27]. Several other species can naturally develop amyloid-like fibrils, including dogs and dolphins [46, 47]. However, the content and distribution of these fibrils across body tissues largely differ from what is found in clinical cases of AD. Within the aging human brain, the initial manifestation of Aβ deposition is especially concentrated in specific areas of the DN, such as the PCu, the PCC, and the orbitofrontal cortex (OFC) [44]. Regions of the DN become less active when engaging in intricate tasks that require substantial attention and conversely become more active when the brain is in its baseline or resting state [48]. While rats and nonhuman primates also have a DN, the functional specialization of the network is believed to largely differ across species [49, 50]. In humans, the DN is thought to map on cognitive functions enabled by conceptual processing such as mind wandering [51], remembering the past [52], envisioning the future [52], and considering the thoughts and perspectives of others [53]. A recent interregional analysis of DN connectivity singled out the fornix fibres among 48 anatomical tracts as most strongly associated with DN gray matter volumes in ~10,000 UK Biobank participants [54]. The fornix, which carries fibre bundle axons from the CA1 and subiculum subregions of the HC, propagates the only hippocampal output signals that directly go to the ventromedial and orbitofrontal cortices of the DN [13, 38]. This pathway is mainly involved in spatial memory and navigation [55, 56]. The HC subiculum, presubiculum, and parasubiculum are also believed to have direct connections to the hypothalamus via the fornix [57]. The hypothalamus plays a crucial role in regulating the release of sex hormones through its control of the pituitary gland. The hypothalamic–pituitary–adrenal (HPA) axis is itself

thought to be regulated by gonadal hormones from early life to adulthood [58]. These connections may serve as a potential pathway, enabling the transmission of negative human experiences that engage DN functions—such as rumination [59, 60], neuroticism [61], and loneliness [62]—to the HPA axis, influencing the stress response and aging process in sex-specific ways. Loneliness, one of the major risk factors for ADRD amongst older adults [63], has indeed been linked to microstructural covariation between the HC and DN [62]. However, the relationship between loneliness and ADRD may vary between sexes. Existing evidence suggests that males experiencing loneliness may have a higher likelihood of developing dementia compared to females [64]. The connection between the DN, the HC, and the HPA axis holds the potential to magnify sex biases in the processing of adverse life events that engage DN functions, such as loneliness. This, in turn, may contribute to distinct patterns of ADRD susceptibility in males and females.

Several environmental, lifestyle and personality-related factors have been shown to either amplify or mitigate the impact of *APOE* alleles. Environmental pollution is believed to precipitate signs of cognitive deficits and amyloid pathology in ε4 carriers, as early as in childhood and adolescence [65, 66]. In a similar vein, engaging in physical activity could decrease the risk of developing dementia amongst ε4 noncarriers, while no such association was observed amongst ε4 carriers [67]. With regards to personality traits, recent evidence has shown that having a positive outlook on aging, such as feelings of usefulness, can amplify the protective effect of *APOE* ε2 against cognitive decline [68]. In this population of older adults, positive beliefs on aging seemed to potentiate the protectiveness of *APOE* ε2 against cognitive decline, whereas the presence of negative beliefs was harmful to the extent that ε2 carriers no longer held a significant advantage against ε4 carriers [68]. Yet, none of these previous studies systematically assessed sex differences in the interaction of *APOE* with modifiable risk factors. The extent to which males and females carrying ε4 or ε2 can equally benefit from improvement in underlying health and lifestyle determinants is still the subject of investigation. A better understanding of the intersectionality between genetics, sex, and lifestyle factors may lead to personalized preventive strategies for ADRD, tailored to an individual's unique genetic and environmental profile.

By leveraging the power of machine learning and big data analytics, we aim to provide an unbiased, data-driven account of the phenotypic variability of familial AD risk as a function of sex

in two large cohorts of asymptomatic older adults. In a first step, we will capitalize on the largest uniformly collected epidemiological brain imaging cohort to date, the UK Biobank, to ascribe sex-specific profiles of AD susceptibility linked to microanatomical defined subregions of the HC and DN, two brain systems known for showing early alteration along the AD continuum [12, 69, 70]. Capitalizing on ~1,000 carefully curated phenotypes, we aim to offer a thorough overview of the preclinical manifestation of ADRD as a function of sex, *APOE* polymorphism, and HC-DN co-variation. In a second step, we aim to carry over our findings to a cohort of at-risk participants that was specially designed to track dementia progression in children of Alzheimer's patients: Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD (PREVENT-AD [71]). In addition to externally validating our findings, this second cohort allows us to zoom in on cross-generational sex differences in AD heritability. Precisely, we aim to address how both the sex of the at-risk children and of their parent diagnosed with AD influence the preclinical manifestation of the disease across well-established risk domains (e.g., genetics, cognition, cardiovascular health, blood and cerebrospinal fluid biochemistry, neurosensory assessments, and lifestyle risk factors). Addressing sex biases in AD risk at a population level is crucial for deriving a better understanding of the interplay between genetics and modifiable risk factors. We believe this work paves the way for personalized preventive strategies in AD prevention, where sex is not just a confounding variable, but rather a central factor in shaping targeted interventions, risk evaluation, and individualized therapeutic approaches based on distinct neurobiological characteristics and lifestyle susceptibilities.

Chapter 1: APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementias

APOE and hippocampus-default network co-variation

Chloé Savignac[1], Sylvia Villeneuve[2,3,4,5], AmanPreet Badhwar[6,7], Karin Saltoun[1], Kimia Shafighi[1], Chris Zajner[1], Vaibhav Sharma[1], 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

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 co-dependent 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 dementias (ADRD). By leveraging the power of the ~40,000 UK Biobank cohort, we assessed impacts from the protective *APOE* ε2 and the deleterious *APOE* ε4 Alzheimer's disease alleles on these structural relationships. We demonstrate ε2 and ε4 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* ε2/2 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* ε2 dosage and sex and link *APOE* alleles to inter-individual brain structural differences indicative of ADRD familial risk.

Introduction

Around the globe, >50 million people are living with dementia – a global burden of >1 trillion USD$ annually [1]. By 2050, an estimated threefold 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 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 unpreceded need to characterize 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 two 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 mobilized 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 ε4 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 ε4 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 ε4 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 ε4 carriers hard to replicate [11]. Moreover, because of the overwhelming singular focus on ε4 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* ε2 and ε4 on DN and HC integrity could be crucial in guiding potential treatment avenues, given the ε2-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 hippocampus formation is known for subfield-specific vulnerability to ADRD, at least since the late 1990s [13]. Yet, the hippocampus 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 hippocampus 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 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 hippocampus as a single unit. In a randomized clinical trial, baseline hippocampus 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 co-dependence 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. Capitalizing on the rich phenotyping available for 40,000 UK Biobank participants, our study could confront the effects of *APOE* ε2 and ε4 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 has 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. Capitalizing 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 ~40,000 UKB participants [21]. These advances enabled us to describe ADRD-related patterns of structural co-variation in DN subregions, which were in lockstep with 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* ε2/2, which has a prevalence of <1% amongst 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 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 two 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 two sets were most strongly correlated with each other. Pairs of canonical variates, one for the HC and one 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_Suppplement_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 vs. without family history of ADRD. A 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 (vPFC); 45.9% of total divergences), and posterior midlines structures (posterior cingulate cortex (PCC), precuneus (PCu) and retrospenial 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 hippocampus 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) which 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, 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, vmPFC, vlPFC, pre-supplementary motor area (Pre-SMA), and orbitofrontal cortex (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 temporo-parietal 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 & 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 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 which 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 systematize 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* ε2, often suspected to confer protective benefits [33], and ε4, 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* ε2 vs. ε4 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 three 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 supplementary information (S1-7 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 FDR correction, we observed additional phenotypes linked with erythrocytes count for both sexes. The second most dominant sets of associations for mode 1 centered on environmental phenotypes, such as NO_2 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 PM_{10}). 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 NO_2 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 emphasized 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 three modes with the most brain-phenotypic associations across sexes (i.e., modes 1, 3, and 8) is presented in the supplementary information (S8-10 Figs.) and serve as a complement to their respective Miami plots (Figures 1, 2 and 3) (cf. methods). The phenome-wide profiles derived across these three concomitant regimes of HC-DN co-variation emphasized 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., ε2/2, ε2/3, ε3/3, ε2/4, ε3/4, and ε4/4). 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 six new population-wide indices, one 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 which 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 three 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 ε2/2 (e.g., being a full or part-time student and doing unpaid or voluntary work), ε2/3 (e.g., number of full siblings, looking after one's home or family, family relationship satisfaction, and number of people in household), ε3/4 (e.g., number of full siblings), ε4/4 (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. Ɛ3 carriership was characterized by the early branching of socioeconomic determinants as shown on the dendrograms for ε2/3 (e.g., past tobacco smoking frequency, time spent watching television, paid employment, average household income, and the number of vehicles in the household), ε3/4 (e.g., past tobacco smoking frequency, alcohol intake frequency, time spent watching television, and education score), and ε3/3 (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 ε2 carriers that comprised self-reported traits related to neuroticism as shown on the dendrograms for ε2/2 (e.g., irritability, miserableness, mood swings), ε2/3 (e.g., being worried/anxious and easily hurt), and ε2/4 (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 ε2 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 treated as input variables the main effect of the HC or DN pattern expressions, the main effects of the six *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.

Across a comprehensive set of analyses across 63 ADRD risk factors, we identified the strongest interaction effects in homozygote ε2 carriers. Notably, brain-*APOE* interactions accounted for more variance in several modifiable social and cardiovascular risk factors than did the main effects of *APOE* ε2 and ε4. Across both sexes, ε2 homozygotes showed strong interaction with HC and DN patterns for being a full vs. part-time student. Male ε2 homozygotes showed strong interactions with HC and DN pattern expressions for doing unpaid or voluntary work. In parallel, female ε2 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 ε2 homozygotes that were not identifiable in heterozygotes and non-carriers. While we observed no appreciable sex effect for the interaction of *APOE* ε4/4 and HC-DN co-variation expressions, we found defensible sex-specificity for the role of *APOE* ε2/2. More precisely, we showed strong interactions between *APOE* ε2/2 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 ε2 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* ε2/2 on ADRD risk amongst males. For *APOE* ε2/3 and ε3/3 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* ε2/4 and ε3/4 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* ε4/4 increased the odds of ADRD family history by more than 56% in males across the different models investigated. In females, *APOE* ε2/2 status decreased the odds of ADRD family history by 50% on average, while ε2/3 and ε3/3 status led to decreases of approximately 25% and 17%, respectively. In contrast, *APOE* ε3/4 and ε4/4 status lifted the odds of ADRD family history by approximately 35% and 86%, respectively. Among females, *APOE* ε2/4 carriers were associated with dampened ADRD risk relative to *APOE* ε3/4 carriers. The odds of ADRD family history associated with *APOE* ε2/4 were only increased by 24% in females. This ~10% reduction in ADRD risk, uniquely observed amongst females, could be taken to suggest that ε2 can still be protective against ADRD risk in the presence of an ε4 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* ε2/2 status was associated with a 2-fold increase in ADRD risk. It was considerably stronger than the main risk effect conferred by *APOE* ε4/4. 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 ε2 and ε4 on ADRD risk, with demonstrably stronger brain-*APOE* interactions amongst females.

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 pre-selected 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 three major axes of brain-behaviour. These

explained ~13.8%, ~9.6%, and ~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 emphasized 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 emphasized 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 two 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 two PCA models in parallel, one for each random half subset, on the z-scored correlation coefficients matrices (63 phenotypes x 50 canonical variates). We showed the average projection of each Pearson's correlation coefficient on the three 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 supplementary information (S15 Fig.).

We then repeated the identical pattern-learning workflow sex-stratifying in males and females separately. The top three principal components explained ~12.1%, ~9.9%, and ~9.0% of the total variance in males, and ~13.1%, ~9.5%, and ~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 emphasized 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 emphasized different categories of risk factors in males and females. The male-derived component 3 emphasized socioeconomic determinants (e.g., education score and the number of vehicles in the household). In contrast, the female-derived one emphasized 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 emphasized 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 1000 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 two PCA models in parallel, one for each sex, on the z-scored correlation coefficients matrices (62 phenotypes x 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 emphasizing 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 capitalized on the openly available 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 behaviors 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 hippocampus, 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 ε4/4 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 characterization of our population-derived limbic-cortical 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* ε2/2. At the same time, common epidemiological studies that have reported the protective effect of carrying an ε2 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 capitalizing on *APOE* genotyping and structural brain scans from ~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* ε2 and ε4 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 ε2 allele typically acts to protect against late-onset Alzheimer's disease [22, 33] and against Aβ accumulation [37-42]. Aβ accumulation in ε2 carriers could be delayed by 30 to 50 years compared to ε4 carriers, who start showing Aβ positivity in their early 40s [12, 40, 43]. The protective qualities of ε2 status have been noted even in the presence of an ε4 allele [12]. Nonetheless, the sex-specific impact of *APOE,* especially its ε2 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 amongst 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* ε2/2 on ADRD risk was not statistically robust amongst males, even in a sample of ~20,000 participants. In contrast, we demonstrated a spectrum of ε2 and ε4 effects amongst females such that *APOE* ε2/4 was associated with milder ADRD risk than ε3/4, which in turn was associated with milder ADRD risk than ε4/4. Resilience towards cognitive decline generally observed amongst ε2 carriers could arise from relatively higher

baseline APOE steady-state levels in regions including the HC and frontal cortex as compared to ε4 carriers and ε3 homozygotes [44-47]. Isoform-specific effects related to the APOE protein could be further enhanced by microglia-driven homeostatic responses to Aβ accumulation [48, 49]. In fact, ε2 carriers are biologically more efficient at scavenging Aβ [50]. As a result, Aβ positivity in ε2 carriers with normal cognition is generally detected in much older age (~95 years) as in ε4 carriers (40-55 years) [40]. Older ε2 carriers with amyloid pathology are likewise less likely to be diagnosed with dementia than ε3 homozygotes of the same age [51]. Cell proliferation and survival in the HC are thought to be particularly modulated by estrogens [52-54] which 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 ε2 carriers [56]. These previous elements of evidence are in line with our present finding suggesting that the protective effect of *APOE* ε2 on ADRD risk is sex-specific and also unique to particular HC-DN co-variation patterns. Notably, we found that female ε2 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 ε2 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* ε2 against ADRD previously discussed, carrying an ε2 allele has been associated with elevated risks

for cardio- and neurovascular disorders [62-66]. *APOE* ε2 is indeed limited in its ability to mediate the vascular clearance of cholesterol metabolites and triglycerides which 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 ε4 status might be taken to suggest that *APOE* ε2 is driving the relationship between physical activity and cognitive performance [68-72]. Hypothetically, engaging in physical activity could be particularly beneficial to older female ε2 carriers in counteracting the rising risk of neurovascular complications resulting from the combined effect of *APOE* ε2 and decreased estrogen levels. Bringing support for this claim, we have shown specific interactions between HC-DN signatures and *APOE* ε2/2 genotype in explaining variation in physical activity – an effect that we found exclusive to females. The specificity of this effect to ε2 homozygotes is consistent with previous findings that have associated ε2 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 behavior that favors longevity via exercise in female ε2 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]. Our phenome-wide assay tied mode 1 expressions to blood markers (e.g., erythrocytes, hemoglobin, 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 subregion 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]. Such 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]. Autopsy samples from patients with Alzheimer's disease have further shown reduced pericyte coverage in CA1 and PFC (Brodmann area 9/10). These were two subregions in which we showed ADRD-related structural divergences in mode 1, as compared to healthy blood vessels in controls [94]. 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. In doing so, we

extend the alleged role of vascular integrity in protecting the brain from environmental stressors that might precipitate ADRD onset.

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 [95, 96]. 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 [97, 98]. As a result, male mice show up to 4-fold higher rates of oxidative toxicity in astrocytes, neurons, and mitochondria compared to female mice [97, 99]. 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 amongst the earliest sites of Aβ accumulation [29] and consume some of the highest oxygen levels in the entire brain [100]. 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 [101]. Recent evidence suggests that $A\beta_{1-42}$ acts on reactive oxygen species to induce glucose hypometabolism [102]. One could argue that the combined effect of air pollution and amyloid pathology could be particularly detrimental in exacerbating ADRD risk amongst 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 emphasizing 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 [103]. IGF-1 is secreted by neurons and glia and possibly acts as a neurotrophic factor regulating neuroendocrine function in the central nervous system [103]. Subcutaneous injection of IGF-1 has previously been associated with increased neurogenesis in the adult rat brain [104, 105]. 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 two subfields in which neurogenesis has been observed in rodents [104, 106, 107] and primates [108]. 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 mode 1 and 3 — are hence in accordance with experimental findings from animal models.

Building on the knowledge that ADRD and verbal-numerical reasoning share overlap in underlying genetic architecture [109], 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 [109-112], 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 amongst 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 [113]. 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 [114]. 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 [115]. Social disengagement has indeed been associated with the incidence of cognitive decline amongst older adults [116-118]. In contrast, engaging in social activities has been linked with up to a 40% decrease in ADRD risk [68, 116, 119]. While social support has been associated with a dampened stress response [120], loneliness is thought to affect not only neuroendocrine but also immune functions [121, 122]. Volunteering and having student status, two 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 ε2 carriers. Neuroticism, which is intimately related to loneliness [123], could predispose individuals to ADRD by weakening strong social support ties and increasing chronic stress through dysregulation of the hypothalamus-pituitary-adrenal (HPA) axis [121, 124]. In fact, the HC subiculum, presubiculum, and parasubiculum are believed to have direct connections to the hypothalamus via the fornix [125]. 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 [126] and dementia [127-129]. Yet, no effects of ε4 dosage on cognitive

decline have been observed in neurotic individuals in these previous reports [126, 128]. The absence of a relationship between *APOE* and neurotic traits reported by previous studies might arise from restricting analyses to ε4 carriers [126, 128]. Indeed, the combined analysis of ε4 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 [130]. The ramifications of neuroticism for ADRD risk, which might be underscored by *APOE* ε2, 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* ε2 against cognitive decline [131]. The protective effect of *APOE* ε2 on cognition was enhanced for individuals with positive beliefs about aging and reduced for those with negative beliefs to the point where ε2 carriers no longer showed a significant cognitive advantage [131]. 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 ε2 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 two main groups when clustered based on their correlation with HC-DN co-variation regimes for ε2 homozygotes. Our study thus reinforces the detrimental effect of neuroticism on ADRD risk and characterized its unique interplay with HC-DN co-variation expressions in ε2 homozygotes.

Conclusion

In sum, the typically protective benefits conferred by *APOE* ε2 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 characterizing the functional correlates of ε4 carriership. Our present study has reconciled these two approaches by contrasting profiles of brain-behaviours associations characteristic of *APOE* ε2 and ε4 in a large epidemiological cohort of ~40,000 participants. In doing so, we were uniquely positioned to illuminate sex-specific associations with modifiable risk factors that were unique to ε2 and ε4 homozygotes. Key risk factors relevant to ε2 carriers included neuroticism, social disengagement, and physical inactivity. In contrast, environmental phenotypes that repeatedly emerged in our results as being linked to ADRD risk were characteristic of ε4 homozygotes. These distinct risk factors could guide potential clinical interventions and governmental policies.

Declaration of interests

The authors declare no competing financial or non-financial interests.

Methods

Population data source

The UK Biobank (UKB) is a large-scale data-collection initiative that offers in-depth information on ~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 multi-modal imaging scans that are routinely augmented and will extend to ~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, UK [132]. 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 HC and 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 (SNP) genotyping (rs429358 and rs7412) and brain-imaging measures (47% men and 53% women). When recruited, these participants were aged 40-70 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 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 U.S. Federal Government's National Institutes of Health, to characterize 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 (2006-2010). 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 [133-135]. There was 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 [136, 137]. 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', that is, the relatives of affected individuals, could allow for a more complete characterization of disease risk amongst individuals before the onset of clinical symptoms [134]. 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 [134, 135].

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 [132, 138]. Noise was removed utilizing 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 [139] and FLIRT [140, 141], as well as non-linear registration to MNI152 standard space at 1 mm resolution using FNIRT [142]. 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, [143]) to generate full bias-field-corrected images. In turn, SIENAX [144] was used to derive volumetric measures normalized for head sizes.

Parcellation of the DN was anatomically guided by the Schaefer-Yeo reference atlas [145]. 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 (~0.1mm 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 [146, 147], 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, age^2, sex, sex*age, and sex*age^2. 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 two high-dimensional variable sets [23, 148, 149]. The first variable set, X, was constructed from subject-level grey matter volume in DN subregions (number of participants x 91 DN parcels matrix). The second variable set, Y, was constructed from HC subregion volumes (number of participants x 38 HC parcels matrix). The two 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 maximizing the linear correlation between low-rank projections from two variable sets or data matrices [23]. The two sets of linear combinations of the original variables are obtained by optimizing the following target function:

$$L_X = XV \qquad L_Y = YU$$
$$l_{X,l} = Xv_l \qquad l_{Y,l} = Yu_l$$
$$corr(l_{X,l}, l_{Y,l}) \propto l_{X,l}^T l_{Y,l} = max,$$

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 lth column of L_X, and $l_{Y,l}$ is the lth column of L_Y.

Our CCA application thus sought to identify linear combinations of X and Y that optimize 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 maximized 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 maximize the symmetric relationship between the data matrices of DN subregion volumes (X) and HC subregion volumes (Y). In doing so, CCA identified the two concomitant projections Xv_l and Yu_l that optimized 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 two 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, 150], we robustly characterized 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 [151]. 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 two CCA models in parallel by fitting one separate model to each of the two 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 one 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 two 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 [147]. In doing so, we obtained a non-parametric estimate of how ADRD risk is manifested in specific subregions for each of the 25 examined HC-DN signatures.

SNP genotyping: six variants of APOE gene

We capitalized on our large sample size to demystify the HC-DN co-variation expressions associated with ε2 allele and ε4 allele homozygotes compared to their heterozygous counterparts for the ε2, ε3, or ε4 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 two SNPs: rs429358 and rs7412. *APOE* ε4 was determined as the combination of rs429358(C) and rs7412(C). *APOE* ε2 was determined as the combination of rs429358(T) and rs7412(T). *APOE* ε3 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: ε3/3 (59.3%), ε3/4 (23.1%), ε2/3 (12.4%), ε2/4 (2.4%), ε4/4 (2.2%), and ε2/2 (0.6%), which correspond to frequencies expected from a population primarily composed of people from European decent [22]. Contrasting the effect of ε2 vs. ε4 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 ε4 carrier vs. non-carrier status.

Phenome-wide analysis of brain-behaviour associations in relation to ε2 vs. ε4 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 ε2 vs. ε4 allele dosage is manifested in inter-individual expressions of HC-DN co-variation and how these manifestations, in turn, relate to UKB traits amongst 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 ~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 [152], https://github.com/MRCIEU/PHESANT) for further refinement. In addition to data cleaning and normalization, PHESANT categorized the data as belonging to one of four 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 [153, 154] 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 [154], 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 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 ε2 vs. ε4 dosage for a given individual. Investigations of *APOE* ε4 dosage effects have been prevalent in brain imaging research [112, 155, 156].

The Alzheimer's disease research community has widely endorsed encoding ε4 dosage in a stepwise fashion, that is, based on the number of allele copies carried by a given patient [110, 112, 155, 156]. 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 [110]. Lyall and colleagues, however, missed looking at ε2 dosage despite the well-established association between the ε2 allele and neurovascular diseases [62, 63]. More recently, APOE ε4 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 [112]. The authors, however, did not benefit from investigating hippocampus anatomical segmentations besides the standard head/body/tail subdivision [112]. Again, APOE ε2 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 ε2 allele, on average [156].

Consequently, the present study builds on the widely shared belief that the ε2 and ε4 alleles have largely opposing effects on Alzheimer's risk and pathophysiology [38, 157, 158]. We sought an analogous composite dosage scale that readily captures opposite effects in modeling the

hippocampus and DN volume variation dependent on the copy number of ε2 and ε4 alleles. We thus created a bivariate dosage scale by summing up positive 'ε2' and negative 'ε4' alleles, such that a homozygous individual carrying *APOE* ε2/2 would have a score of +2 and one carrying *APOE* ε4/4 a score of -2. The neutral *APOE* ε3 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 ε2 and ε4 in a single model. In doing so, we stayed faithful to our overarching goal of unraveling their adversarial impact on HC-DN co-variation.

Aiming to capture possible sex-specific effects, we regressed the ε2 vs. ε4 dosage on inter-individual expressions of a given mode in males and females separately. We thus estimated 2 * 25 different base models, one for each HC-DN signature and each sex, that each had two 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 ε2 vs. ε4 dosage as a function of HC-DN co-variation expressions. For each subject and mode combination, we asked *what would the expected ε2 vs. ε4 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 ε2 vs. ε4 dosage associated with their individual expression of HC-DN co-variation. For each mode, we selected the 5^{th} and top 95^{th} percentiles to identify the top 5% and lower 5% of individuals who were more vs. less likely to develop ADRD based on the derived ε2 vs. ε4 dosage risk. We focused on the extreme of the dosage distribution to target the brain-*APOE* associations especially linked to ε2 and ε4. The analogous approach is widely adopted in genome-wide analyses to remove associations not directly linked to the target genotype [159, 160].

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, two 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 = 5.11e-5). Second, we used the false discovery rate (FDR), another popular adjustment, although less stringent than Bonferroni's correction. The false discovery rate [161] was set as 5% [138, 162, 163] and computed for each HC-DN signature in accordance with standard practice [164]. For the sake of visualization, we used Miami plots to compare the profiles of brain-behaviour associations derived from males and females. For visualization 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 non-linear 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 six one-hot encoded *APOE* genotypes (i.e., ε2/2, ε2/3, ε3/3, ε2/4, ε3/4, and ε4/4) such that participants without a given genotype were zeroed out. The six ensuing matrices (number of participants x 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 six genotype-specific matrices and the z-scored risk factor matrix (37,291 participants x 63 risk factors) to investigate risk-anatomy links. Spearman's correlation is a nonparametric metric of statistical dependence between the rankings of two variables that can be used to capture monotonic non-linear phenomena. The Spearman's correlation coefficients reduce to the Pearson's correlation between the rank values of two variables and hence range from -1 (inversely proportional association) to +1 (proportional

association). We obtained a new cross-association matrix $X \in \mathrm{R}^{63 \times 50}$ which represented the Spearman's correlation between the 63 risk factors and the 50 canonical variates for each of the six *APOE* genotypes. The obtained Spearman's correlation coefficients thus carried the non-linear association strength of a given risk-anatomy link for a particular *APOE* genotype.

For each of the six *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 [165] 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 minimizing the total within-cluster variance defined as the error sum of squares:

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

where d_{ij} represents the squared Euclidean distance between two 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 six *APOE* genotypes, we could thus create a dendrogram that represented the distance in Euclidian space between the clusters retained after three levels of branching. The level of branching refers to the number of divisions from the final merge. The dendrograms allowed us to visualize the clustering results for each of the six *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 two descendent links across corresponding hierarchical merging steps in the six 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 characterize 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 two regression models: one for its HC pattern and one 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 one HC or DN canonical variate, the six *APOE* genotypes (ε2/2, ε2/3, ε3/3, ε2/4, ε3/4, and ε4/4), and six interaction terms capturing the non-linear association between each of the six *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 * 1000 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 randomized 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 one HC or DN canonical variate, the six *APOE* genotypes (ε2/2, ε2/3, ε3/3, ε2/4, ε3/4, and ε4/4), and six interaction terms capturing the non-linear association between each of the six *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 principal component analysis (PCA) [166]. 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 x 50 canonical variates) and the z-scored risk factor matrix (number of participants x 63 risk factors). We obtained a new matrix $M \in \mathrm{R}^{63 \times 50}$, 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 [167]. More formally, solving the SVD problem took the following form:

$$M = U\,S\,V^{T}\,,$$

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

We retained the top three singular vectors and expressed our correlation matrix in terms of the dot product $US \in R^{63 \times 3}$ to be able to represent the latent-factor projections of M onto the new three-dimensional latent space. In doing so, we obtained the distinct expression levels of the 63 risk factors for each of the top three brain-behaviour association axes (i.e., principal component expressions). These three 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 two 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 three 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 three main axes of brain-behaviour associations derived from each random subset. If the PCA solution is robust, similar groups of risk factors should be emphasized 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 three first principal components for random subsets 1 and 2 across the 1000 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 three 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 emphasized 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 1000 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 three first principal components for male- and female-derived solutions in each of the 1000 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); [168]) 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, standard deviation [SD] 5 years) [168]. Participants of the PREVENT-AD initiative have undergone extensive annual health and cognitive assessments for up to five 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: ε3/3 (51.2%), ε3/4 (33.1%), ε2/3 (10.5%), ε2/4 (3.0%), ε4/4 (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 four 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, age^2, sex, sex*age, and sex*age^2 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.

Data and code availability

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 supplementary figures have been publicly deposited online (DOI: 10.5281/zenodo.7126809).

Figures

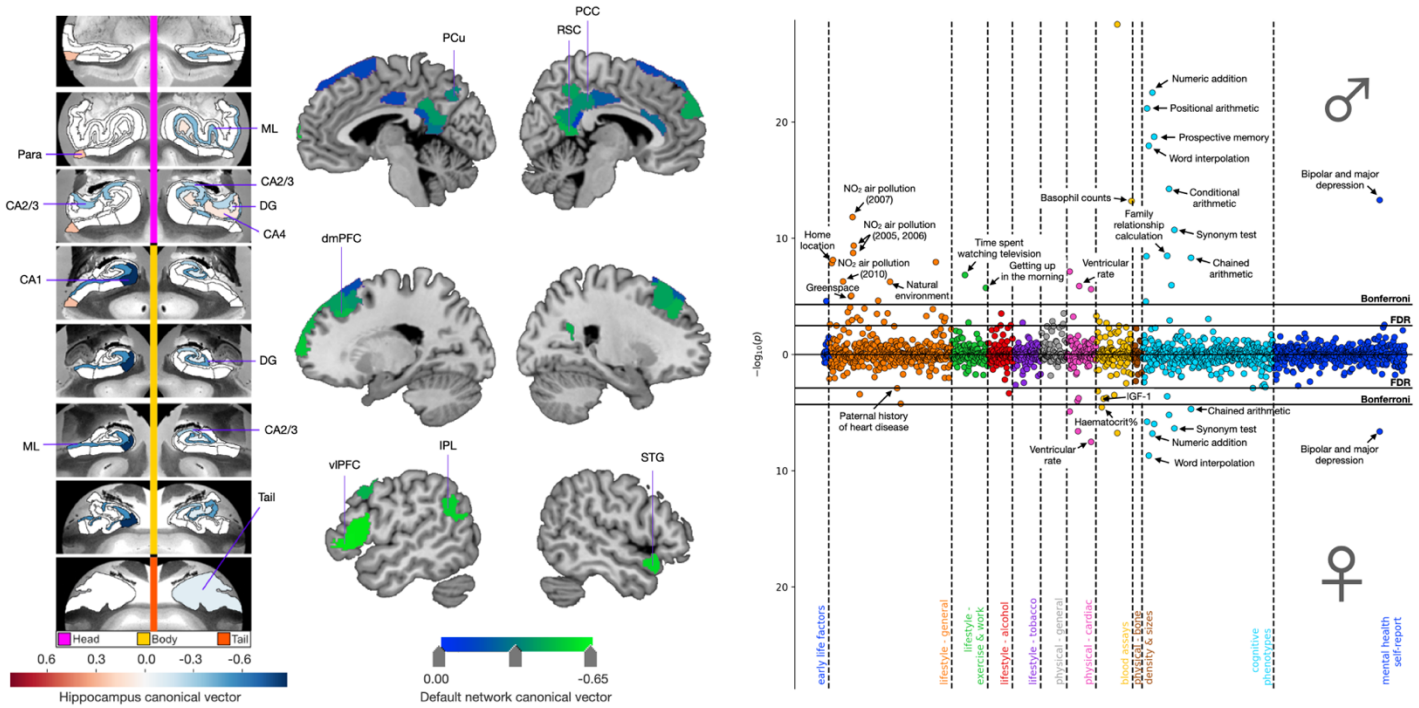


Figure 1. Cognitive, environmental, and cardiovascular phenotypes show sex-specific associations with *APOE* dosage in the context of mode 1. The leftmost and central panels display structural divergences in the HC and DN, respectively, on mode 1 for the group difference analysis of ADRD family history. We identified 12 HC hits, mostly located in the cornu amonis (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 & diastolic blood pressure, insulin-like growth factor 1 (IGF-1), and urea) that extended beyond physical traits shared with males (e.g., cardio-respiratory fitness, and ventricular & 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). ML = molecular layer, Para = parasubiculum, DG = granule cell layer of the dentate gyrus, PCu = precuneus, RSC = retrosplenial cortex, PCC = posterior cingulate cortex, dmPFC = dorsomedial prefrontal cortex, vlPFC = ventromedial prefrontal cortex, IPL = inferior parietal lobule, STG = superior temporal gyrus, FDR = false discovery rate correction.

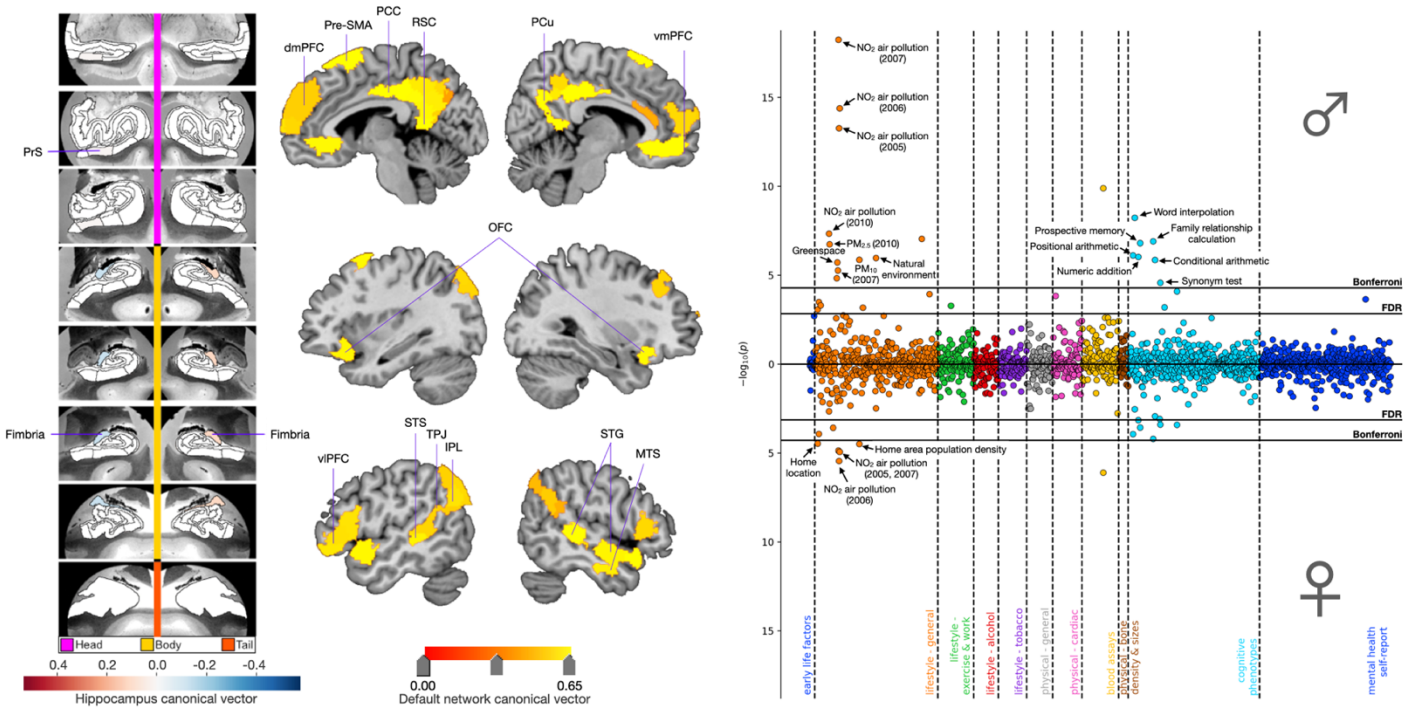


Figure 2. *APOE*-modulated associations for mode 3 revealed a prominence of cognitive and environmental phenotypes in males. Shown here are ADRD-related subregion divergences for mode 3 for the HC (leftmost panel) and DN (central panel). We identified focalized hits to the fimbria and presubiculum with 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 sub-questions of 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). PrS = presubiculum, dmPFC = dorsomedial prefrontal cortex, Pre-SMA = pre-supplementary motor area, PCC = posterior cingulate cortex, RSC = retrosplenial cortex, PCu = precuneus, vmPFC = ventromedial prefrontal cortex, OFC = orbitofrontal cortex, vlPFC = ventrolateral prefrontal cortex, STS = superior temporal sulcus, TPJ = temporo-parietal junction, IPL = inferior parietal lobe, STG = superior temporal gyrus, MTS = middle temporal sulcus, FDR = false discovery rate correction.

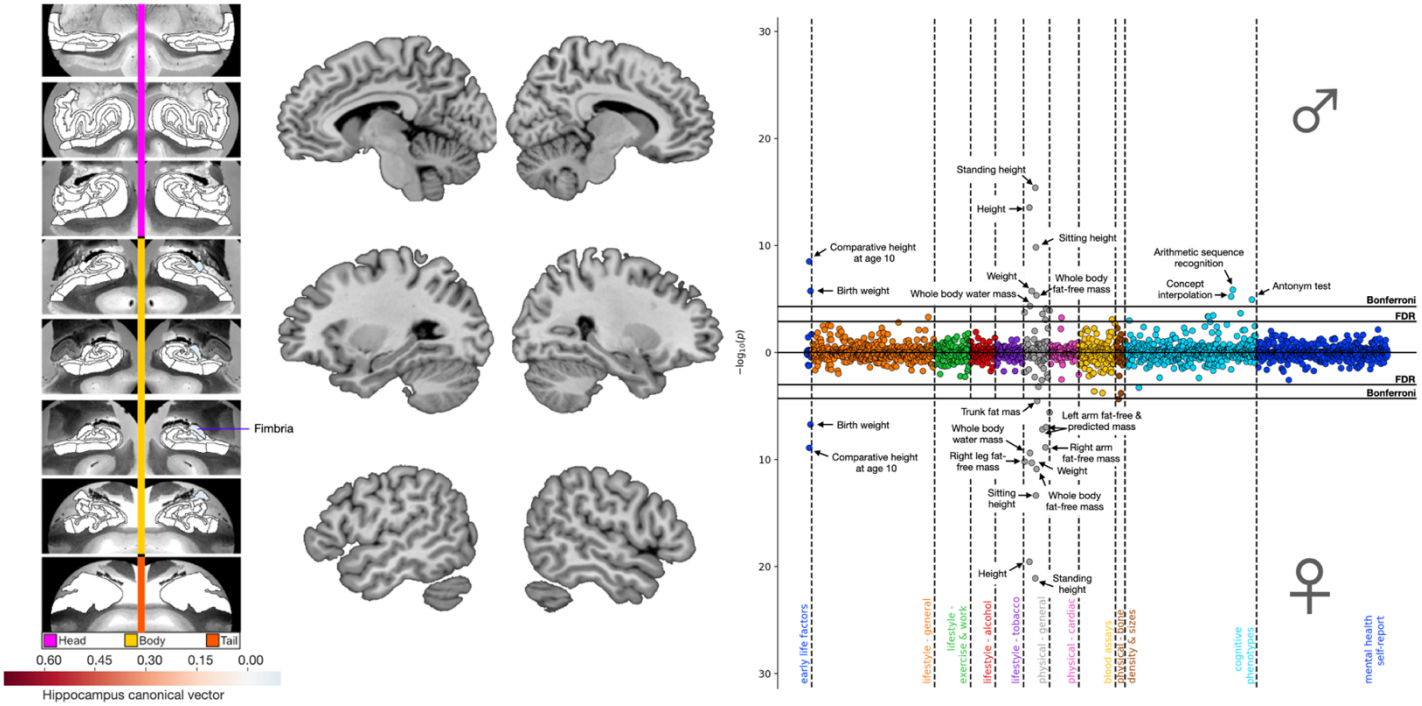


Figure 3. *APOE*-modulated associations for mode 8 linked lipid metabolism to deviation of the fimbria. Shown here are ADRD-related subregion divergences for mode 8 for the HC (leftmost panel) and DN (central panel). We identified a focalized divergence to the fimbria with no corresponding DN hits. In 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 sub-questions from 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). FDR = false discovery rate correction.

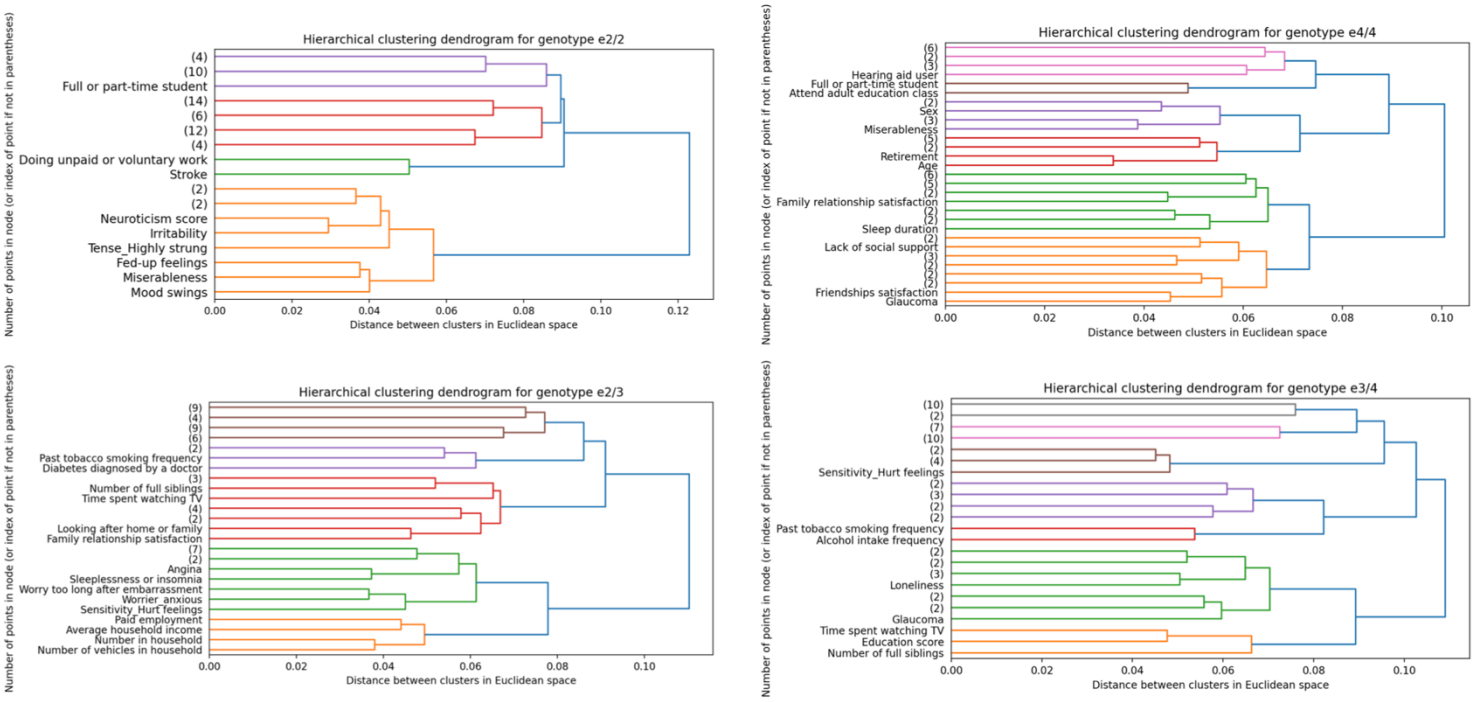


Figure 4. Neuroticism-related phenotypes show unique risk-anatomy links in ε2 carriers. To test for risk-anatomy links, we computed the Spearman's correlations between the population-wide HC and DN co-variation patterns, multiplied by each APOE of the six genotypes and the 63 pre-selected Alzheimer's disease risk factors. 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 which indicate the distance between each cluster identified when retaining three levels of branching for APOE ε2/2 (upper left; N=217), ε2/3 (lower left; N=4,625), ε4/4 (upper right; N=822), ε3/4 (lower right; N=8,613). The dendrograms for ε3/3 and ε2/4 can be found in the supplementary 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. Ɛ3 carriership was characterized 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 ε2/3, ε3/4, and ε3/3 (S12 Fig.). While clusters of social engagement and socioeconomic determinants were shared across different APOE genotypes, we found that neuroticism was uniquely associated with ε2 carriership. Indeed, the dendrogram for ε2/2, ε2/3, and ε2/4 (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 ε2 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)

Figure 5. Brain-*APOE* ε2/2 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 pre-selected 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* ε2/2 interacts with HC and DN co-variations patterns in estimating being a full or part-time student, with stronger coefficients observed for males on the HC side (regression models 1-25). Similarly, on the middle panels, we show that *APOE* ε2/2 preferentially interact with HC and DN canonical variates in estimating doing unpaid or voluntary work in males. In the bottom panels, we show that *APOE* ε2/2 interact with selective HC and DN canonical variates in estimating engaging in strenuous sport in females. We have thus shown that *APOE* ε2/2 interacts preferentially with HC-DN co-variation patterns in estimating social lifestyle in males and physical activity in females. These interactions profiles suggest that ε2, and not ε4, is driving most of the brain-genes 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).

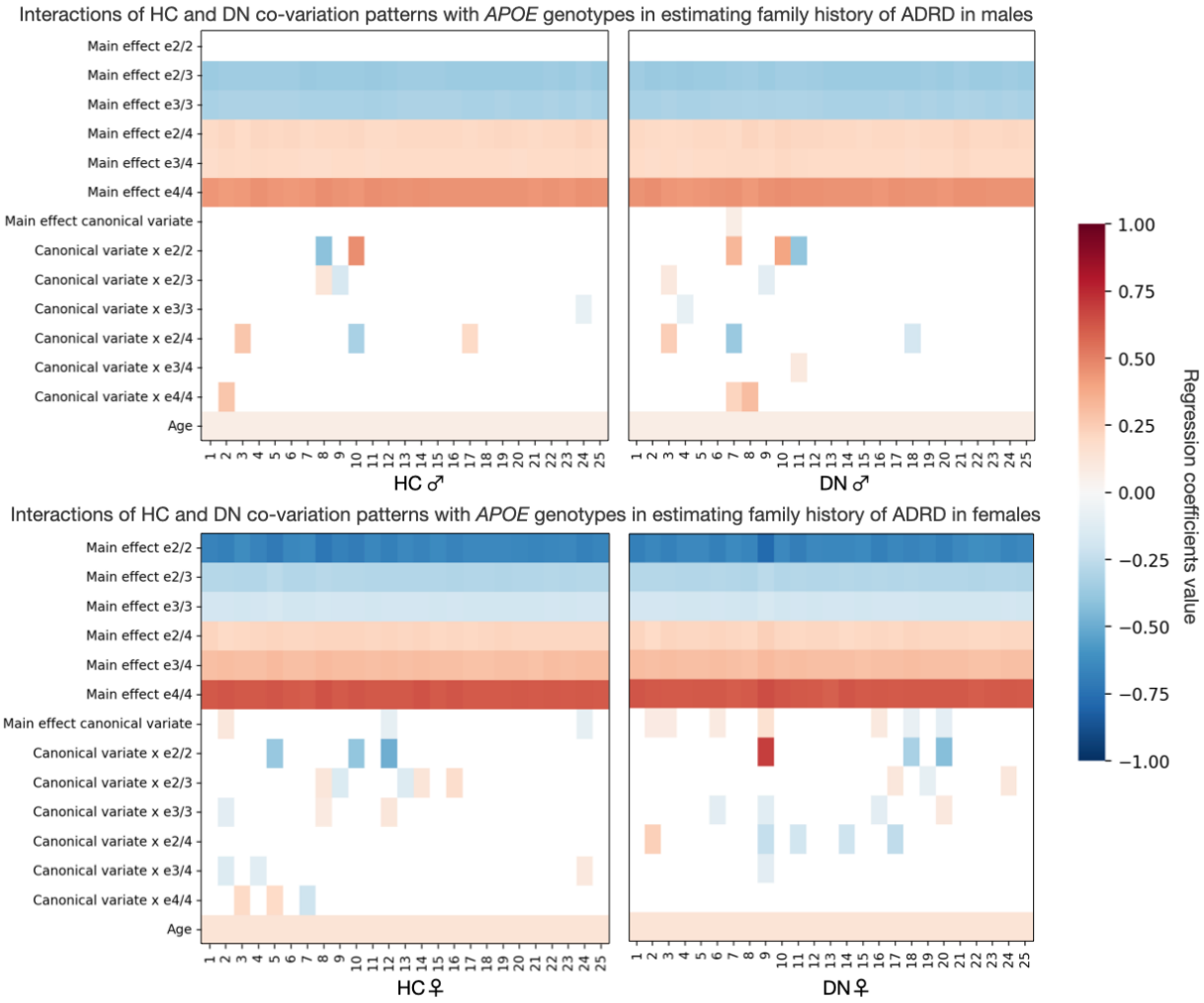


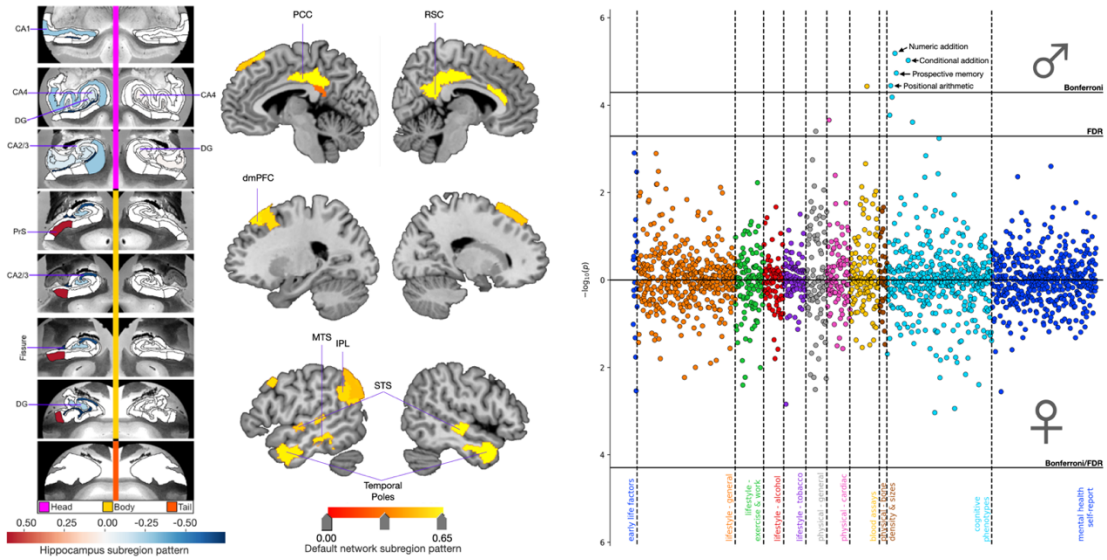
Figure 6. The protectiveness of ε2 is sex-dependent and modulated by HC-DN co-variation patterns. In separate analyses for males and females, we tested whether HC-DN signatures interacted with *APOE* genotypes in explaining variance in family history of ADRD. Separate analyses were run for males (higher 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* ε2/2 against ADRD risk was only statistically robust in females. We also showed a spectrum in the opposing effects of ε2 and ε4 amongst females, such that ε2/4 was associated with a lower increase in ADRD risk than did *APOE* ε3/4, which in turn was associated with lesser risk than ε4/4. We further found that the protectiveness of *APOE* ε2/2 interacts with brain structure and can even lead to an increase in ADRD risk amongst females with a strong expression of mode 9. These interactions profiles suggest that the protectiveness of ε2/2 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).

Tables

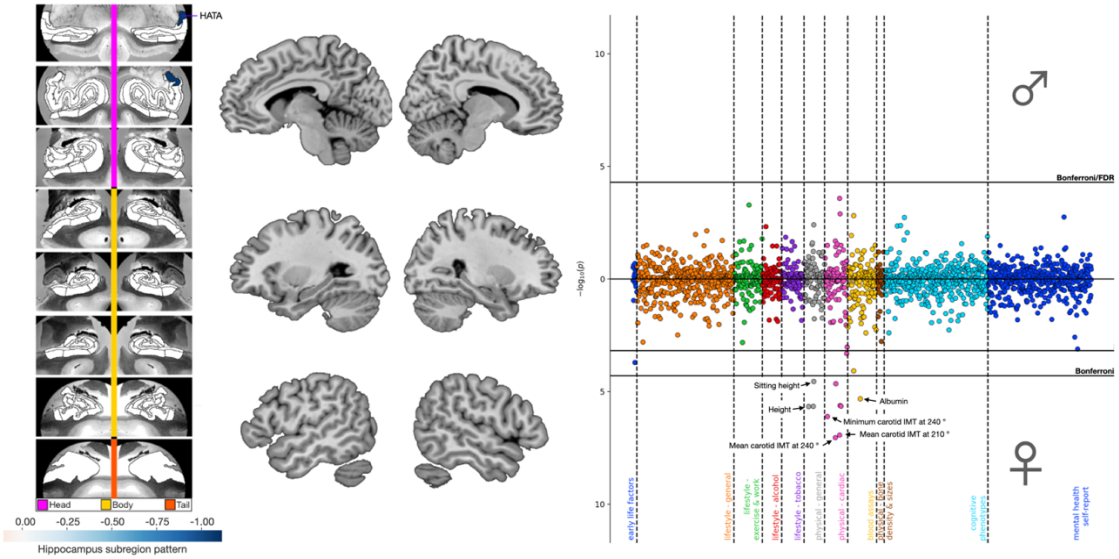
	ε3/3	ε3/4	ε2/3	ε2/4	ε4/4	ε2/2
N (%)	22,129 (59.3)	8,613 (23.1)	4,625 (12.4)	885 (2.4)	822 (2.2)	217 (0.6)
Age, Mean ± SD	54.9 ± 7.5	54.5 ± 7.4	55.0 ± 7.5	55.0 ± 7.5	54.3 ± 7.3	54.6 ± 7.5
Sex, *n* (%)						
Females	11,579 (52.3)	4,634 (53.8)	2,464 (53.3)	489 (55.3)	447 (54.4)	117 (53.9)
Males	10,550 (47.7)	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

Table 1: UK Biobank demographic information. Distribution of the demographic information from the UK Biobank participants included in the present study grouped per *APOE* genotypes.

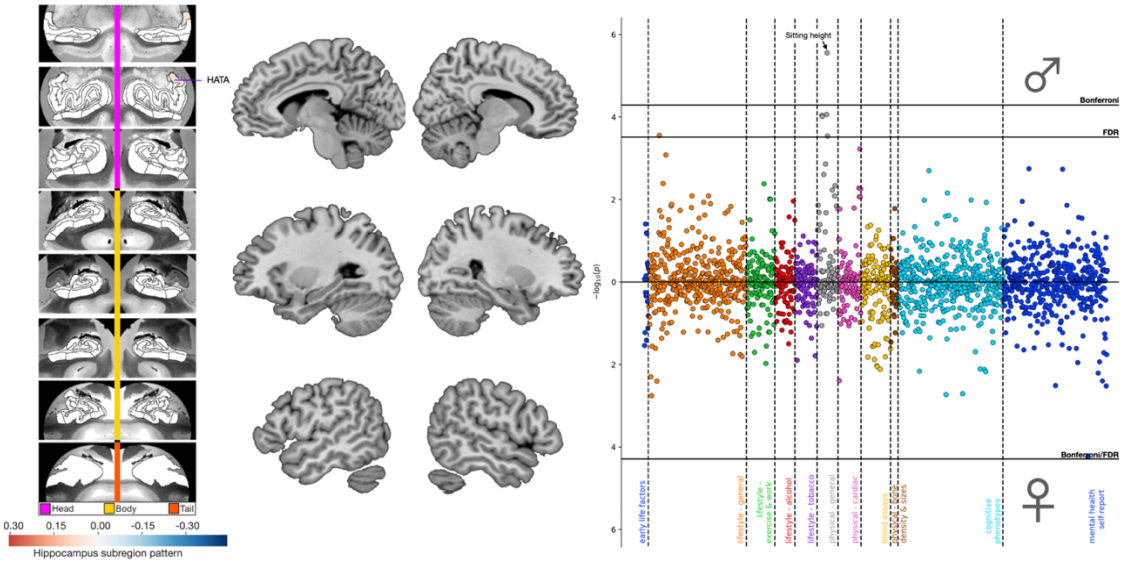
Supplementary Material



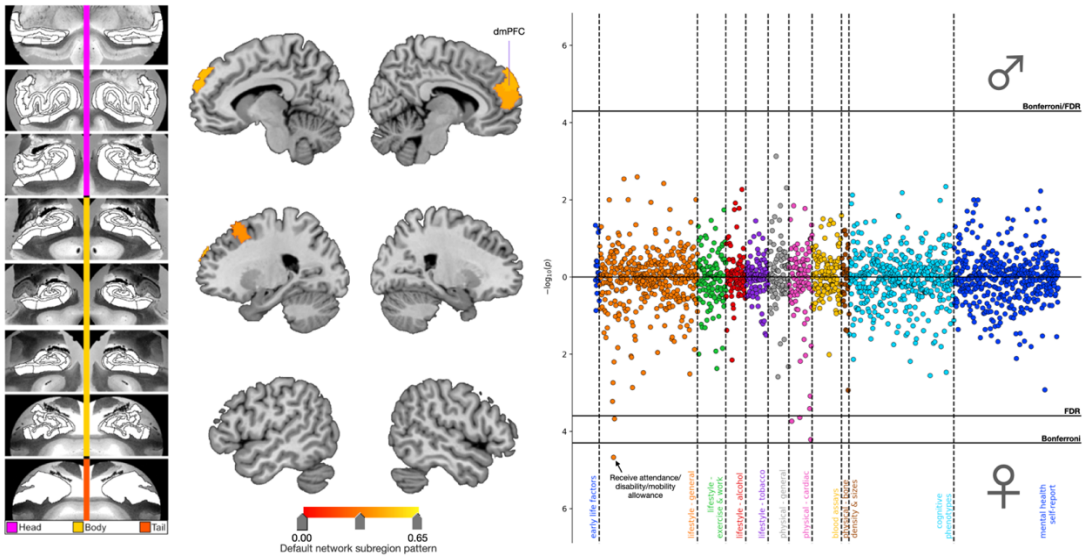
S1 Fig. ADRD-related divergences in HC and DN subregions for mode 2 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 2 for the HC (leftmost panel) and DN (central panel). We identified 10 HC hits, most of them located in the left hemisphere. The strongest 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 restrosplenial 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 sub-questions from 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). CA = cornu amonis, DG = granule cell layer of the dentate gyrus, PrS= presubiculum, PCC = posterior cingulate cortex, RSC = retrosplenial cortex, dmPFC = dorsomedial prefrontal cortex, IPL = inferior parietal lobule, MTS = middle temporal sulcus, and STS = superior temporal sulcus, FDR = false discovery rate correction.



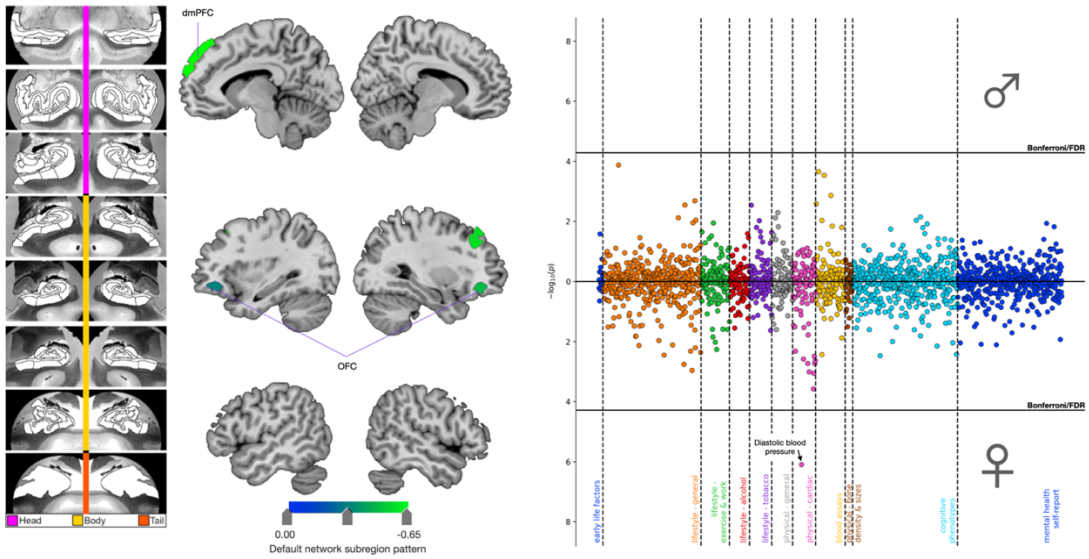
S2 Fig. ADRD-related divergences in HC and DN subregions for mode 6 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 6 for the HC (leftmost panel) and DN (central panel). We identified 1 HC hit to the hippocampus-amygdala transition area with no concurrent 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). HATA = hippocampus-amygdala transition area, IMT = intima-medial thickness, FDR = false discovery rate correction.



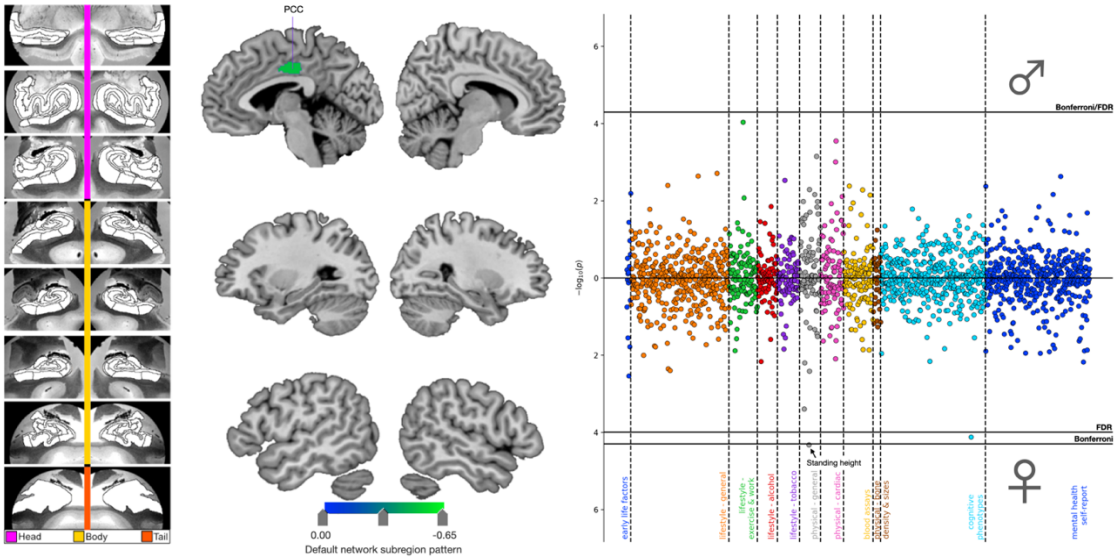
S3 Fig. ADRD-related divergences in HC and DN subregions for mode 10 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 10 for the HC (leftmost panel) and DN (central panel). We identified 1 HC hit to the hippocampus-amygdala transition area with no concurrent 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 one 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). HATA = hippocampus-amygdala transition area, FDR = false discovery rate correction.



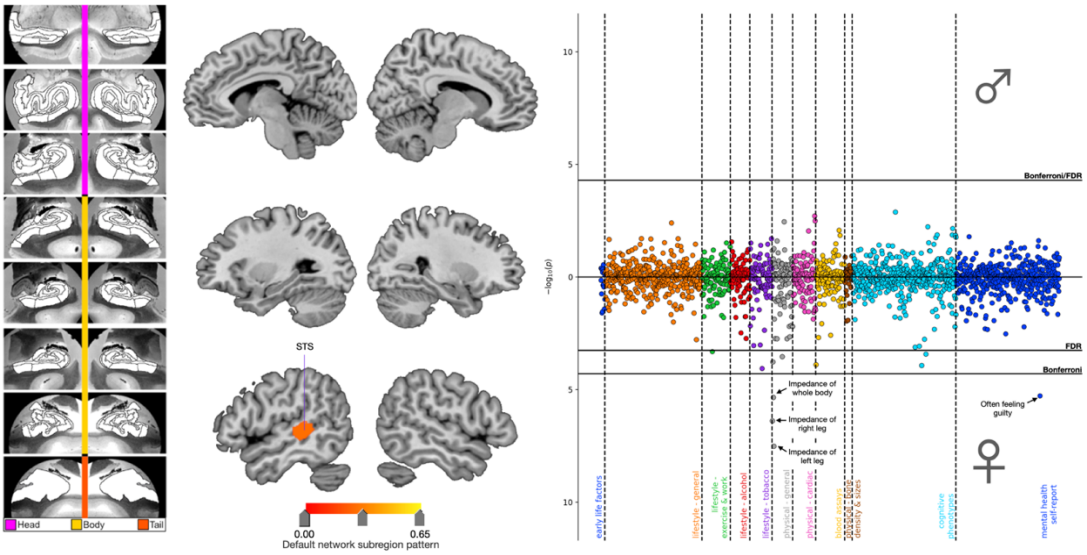
S4 Fig. ADRD-related divergences in HC and DN subregions for mode 4 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 4 for the HC (leftmost panel) and DN (central panel). We identified 4 DN hits to the dorsomedial prefrontal cortex with no concurrent 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 one 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). dmPFC = dorsomedial prefrontal cortex, FDR = false discovery rate correction.



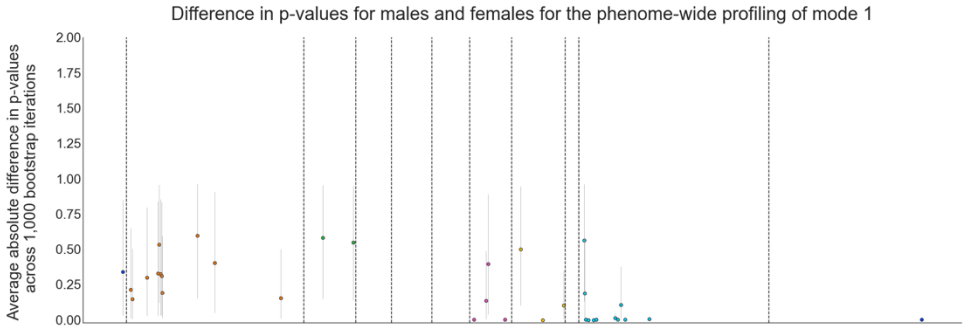
S5 Fig. ADRD-related divergences in HC and DN subregions for mode 7 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 7 for the HC (leftmost panel) and DN (central panel). We identified 9 DN hits to the frontal lobe with no concurrent HC divergences. 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 one 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). dmPFC = dorsomedial prefrontal cortex, OFC = orbitofrontal cortex, FDR = false discovery rate correction.



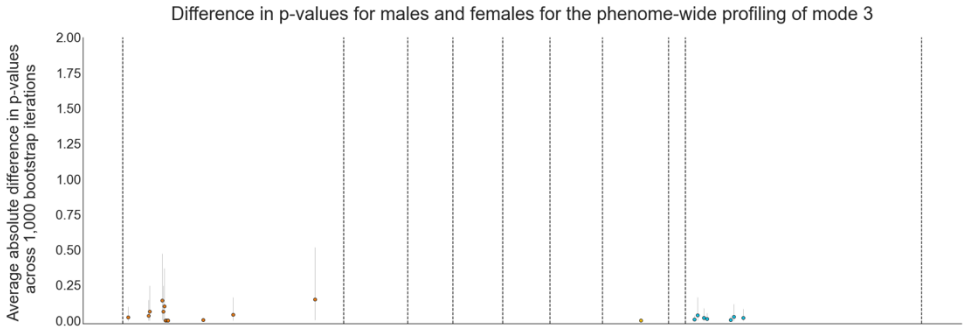
S6 Fig. ADRD-related divergences in HC and DN subregions for mode 11 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 11 for the HC (leftmost panel) and DN (central panel). We identified 1 DN hit to the posterior cingulate cortex with no concurrent HC 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 one 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). PCC = posterior cingulate cortex, FDR = false discovery rate correction.



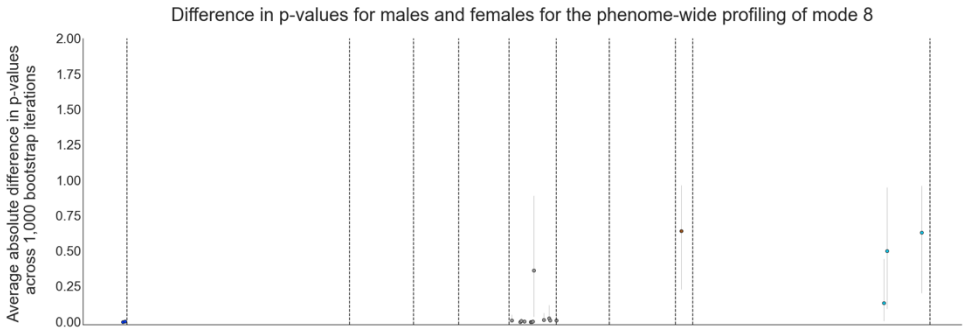
S7 Fig. ADRD-related divergences in HC and DN subregions for mode 13 and the associated phenome-wide profile. Shown here are ADRD-related subregion divergences for mode 13 for the HC (leftmost panel) and DN (central panel). We identified 1 DN hit to the superior temporal sulcus with no concurrent HC 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). STS = superior temporal sulcus, FDR = false discovery rate correction.



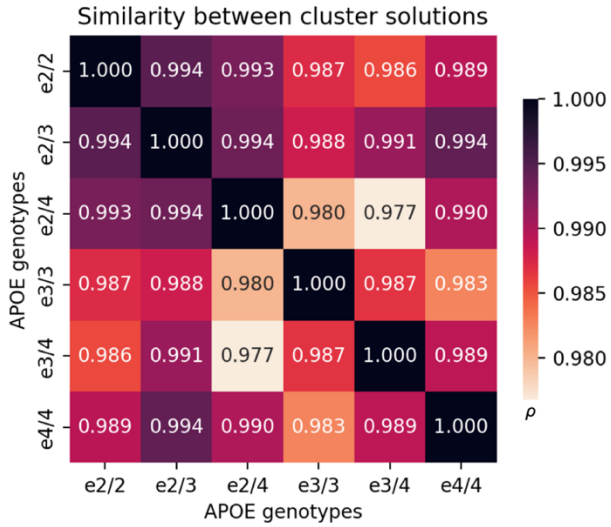
S8 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 1. Absolute difference in p-values for the 33 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling 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).



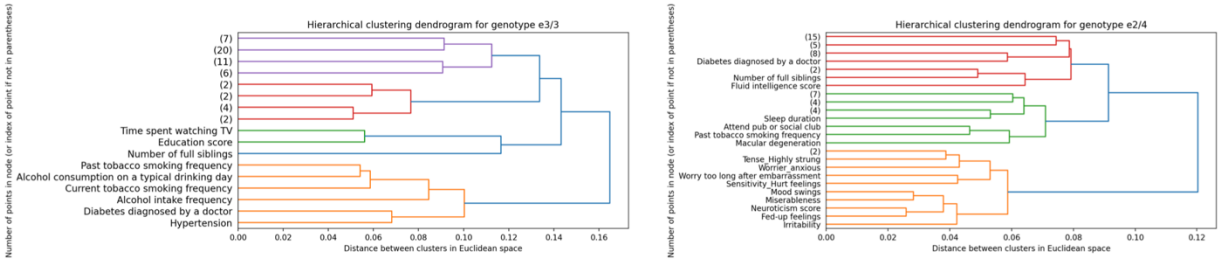
S9 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 3. Absolute difference in p-values for the 20 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling 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).



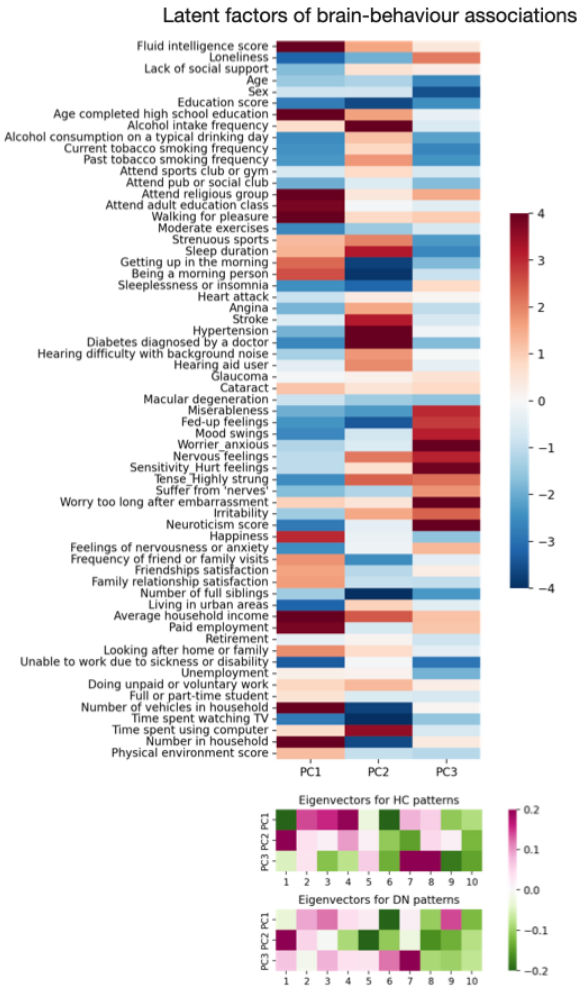
S10 Fig. Difference in associations between males and females for the phenome-wide profiling of mode 8. Absolute difference in p-values for the 18 brain-phenotype associations that passed the Bonferroni correction for multiple comparisons in either males or females in the original phenome-wide profiling 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).



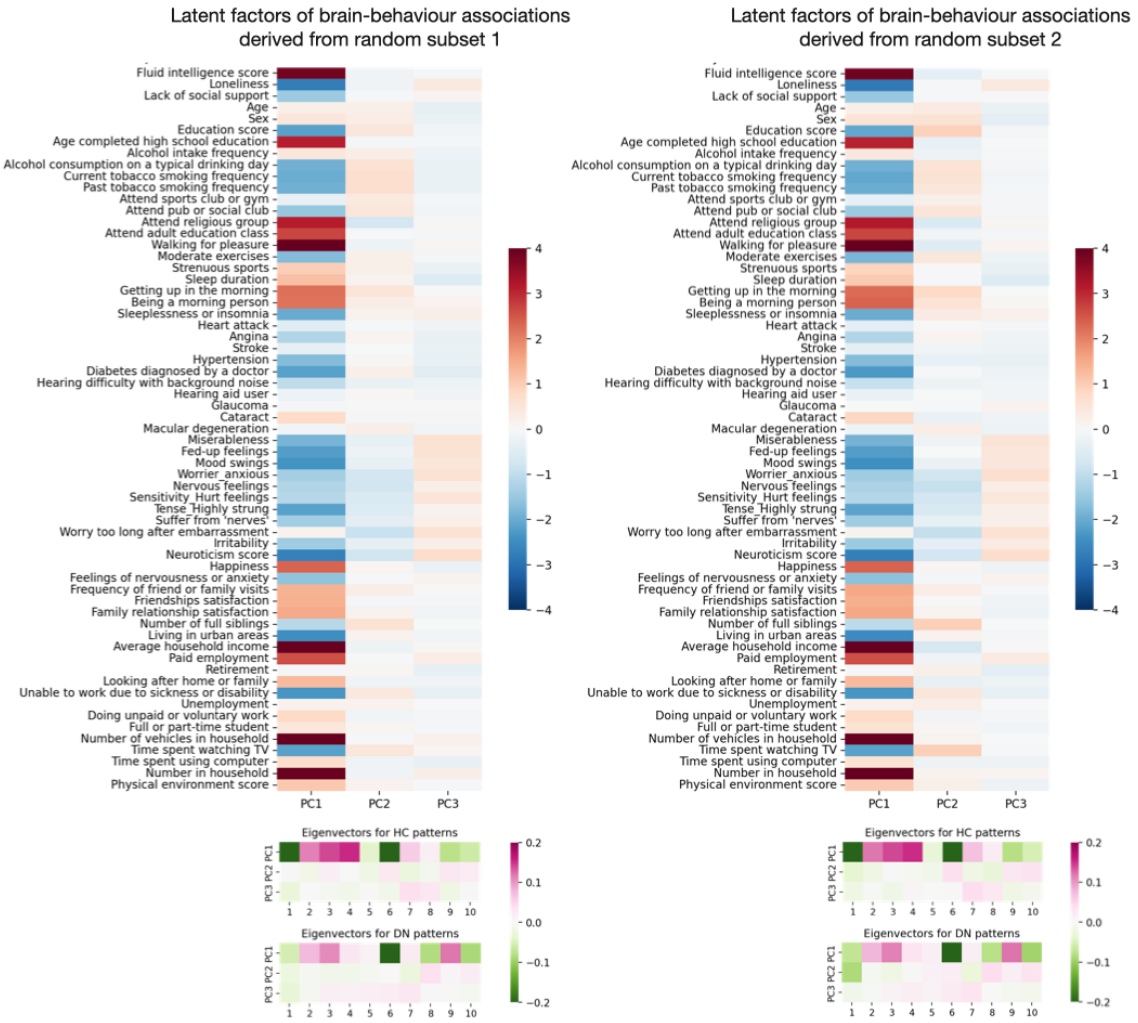
S11 Fig. Similarity between the six genotype-specific clustering models. We computed Pearson's correlation of the distance between the two descendent links of corresponding hierarchical merging steps among the cluster analyses for the six APOE genotypes (i.e., ε2/2, ε2/3, ε3/3, ε2/4, ε3/4, and ε4/4). These derived distances made it possible to formally compare the cluster nodes of analogous dendrograms for each genotype-specific cluster model. We show that ε2 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 heatmap. The most dissimilar cluster models were ε2/4 and ε3/4, followed by ε2/4 and ε3/3, and lastly by ε3/3 and ε4/4. 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).



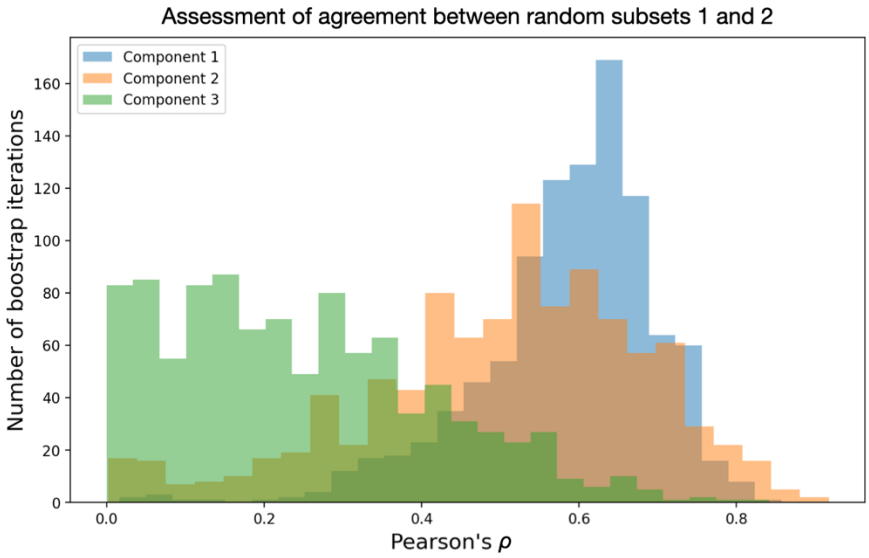
S12 Fig. Ɛ3 carriership shows risk-anatomy links with socioeconomic determinants, while ε2 carriership is associated with neuroticism. We multiplied the population-wide HC and DN co-variation patterns by APOE genotypes ε3/3 (N=22,129) and ε2/4 (N=885) such that participants who do not carry a given genotype were zeroed out. We then computed the Spearman's correlations between these two new vectors and the 63 pre-selected 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 three levels of branching for APOE ε3/3 (leftmost panel) and ε2/4 (rightmost panel). We found the early branching of socioeconomic determinants ε3/3 (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 ε3/3. For ε2/4, 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 ε2 carriers. We thus confirm the association between ε3 carriership and socioeconomic determinants and between ε2 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).



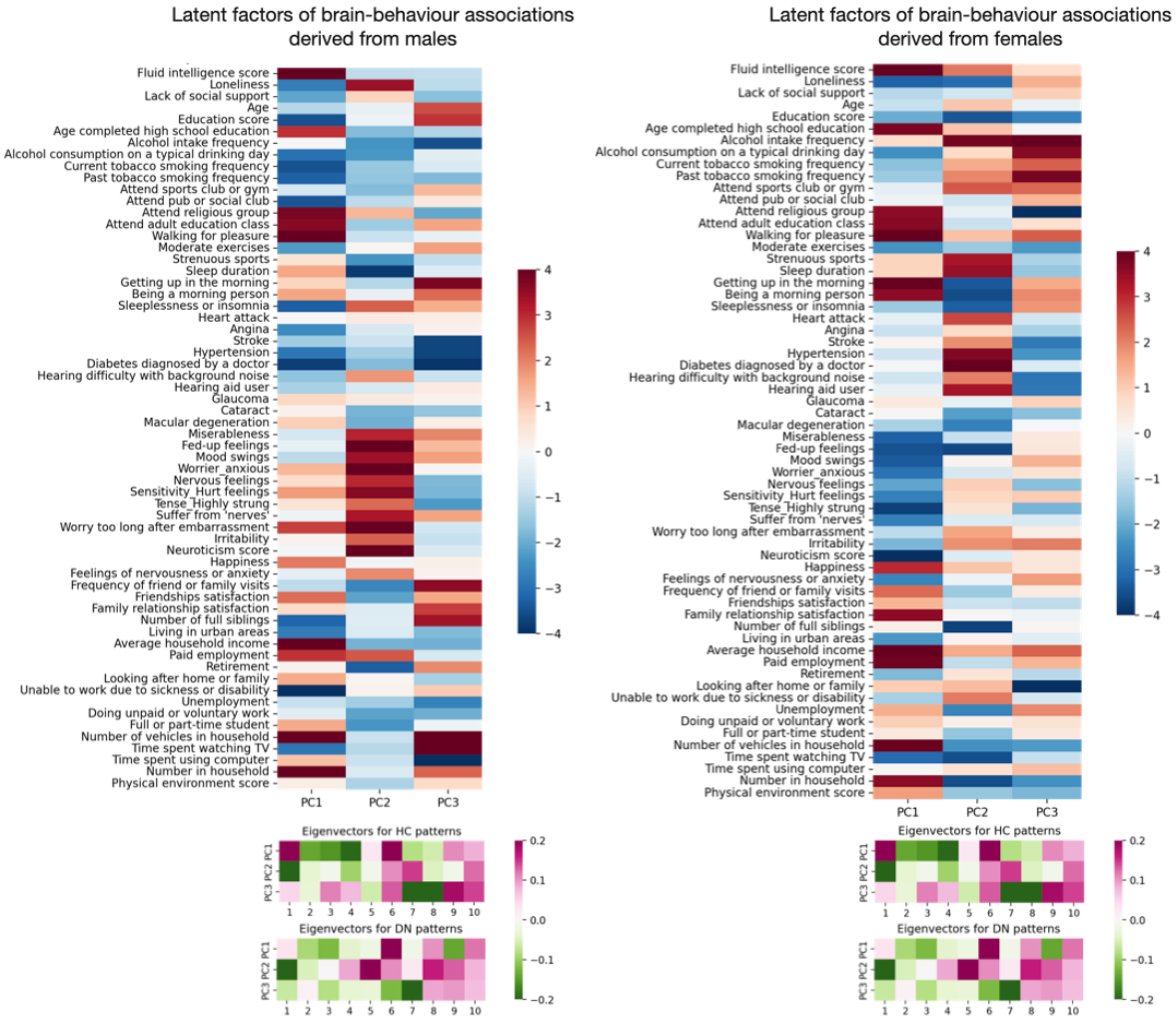
S13 Fig. Latent factors of brain-behaviour associations emphasize satisfaction with social relationships, socioeconomic status, and neuroticism-related traits. We conducted an exploratory principal component analysis (PCA) to disentangle latent factor of brain-behaviour association in our UK Biobank 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 pre-selected 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 principal components (PCs) that explained ~13.8%, ~9.6%, and ~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 three main axes of brain-behaviour associations. The lower plot displays the eigenvectors for the top ten HC and DN co-variation patterns. The first axis of brain-behaviour associations emphasizes 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).



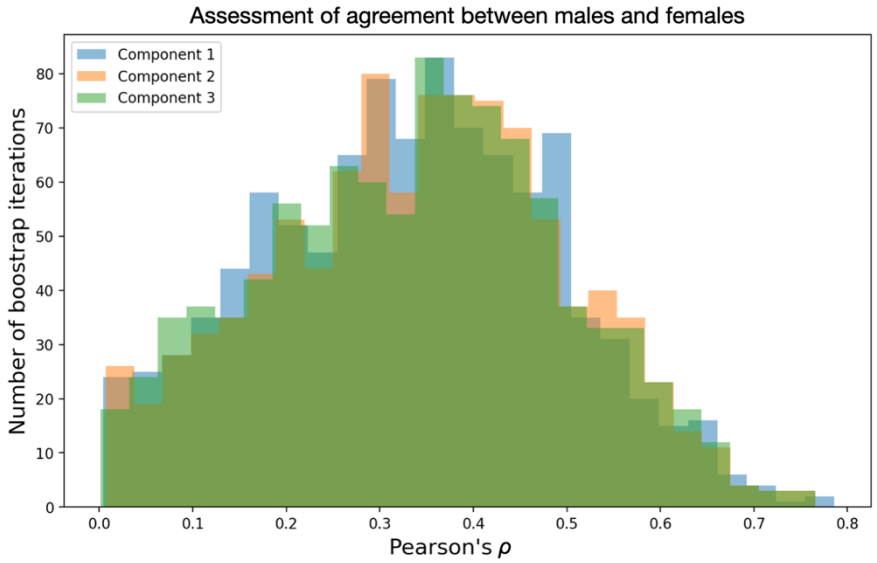
S14 Fig. Reliability assessment of the principal component solution. We assessed the robustness of the derived brain-behaviour association axes by performing a split-half reliability assessment of our principal component solution across 1,000 bootstrap iterations. At each iteration, we drew 37,291 participants with replacements to simulate random participant samples that we could have pulled from the same population. We then derived two 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 two 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 three 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) emphasized. Data underlying this figure can be found at https://github.com/dblabs-mcgill-mila/HCDMNCOV_AD/blob/master/PCA (DOI: 10.5281/zenodo.7126809).



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 three first principal components for random subsets 1 and 2 across 1000 bootstrap iterations. The weights of the first two 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).



S16 Fig. Neuroticism-related items expressed distinctive brain-behaviour associations in males and females. We repeated the principal component analysis 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 pre-selected 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 principal components (PCs). The PCs obtained from males had explained variance of ~14.6%, ~11.9%, and ~9.6%, respectively. The PCs obtained from females had explained variance of ~14.6%, ~11.9%, and ~7.4%, respectively. The upper plots display the projections of the Pearson's correlations onto each of the three axes of brain-behaviour associations for the two sexes. The lower plots display the eigenvectors for the top ten HC and DN co-variation patterns. The projections of the Pearson's correlations onto the two first axes of brain-behaviour association were roughly the same in males and females. In contrast, neuroticism-related items were only emphasized 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).



S17 Fig. Statistical agreement between the PCA solutions for males and females. We computed the Pearson's correlation between the weights of the first three principal components for the sex-specific PCA solutions across 1000 bootstrap iterations. We observed a low agreement between the male- and female-derived PCA solutions on all three 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).

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 ε4/4 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).

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.* The Lancet Public Health, 2022.
3. Livingston, G., et al., *Dementia prevention, intervention, and care: 2020 report of the Lancet Commission.* The Lancet, 2020. **396**(10248): p. 413-446.
4. Poirier, J., et al., *Apolipoprotein E polymorphism and Alzheimer's disease.* The Lancet, 1993. **342**(8873): p. 697-699.
5. Herholz, K., et al., *Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET.* Neuroimage, 2002. **17**(1): p. 302-316.
6. Badhwar, A., et al., *Resting-state network dysfunction in Alzheimer's disease: a systematic review and meta-analysis.* Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 2017. **8**: p. 73-85.
7. Song, H., et al., *APOE effects on default mode network in Chinese cognitive normal elderly: relationship with clinical cognitive performance.* PLoS One, 2015. **10**(7): p. e0133179.
8. Goveas, J.S., et al., *Functional network endophenotypes unravel the effects of apolipoprotein E epsilon 4 in middle-aged adults.* PloS one, 2013. **8**(2): p. e55902.
9. Westlye, E.T., et al., *Increased Hippocampal Default Mode Synchronization during Rest in Middle-Aged and Elderly APOE ε4 Carriers: Relationships with Memory Performance.* The Journal of Neuroscience, 2011. **31**(21): p. 7775-7783.
10. Filippini, N., et al., *Distinct patterns of brain activity in young carriers of the APOE-ε4 allele.* Proceedings of the National Academy of Sciences, 2009. **106**(17): p. 7209-7214.
11. Mentink, L.J., et al., *Functional co-activation of the default mode network in APOE ε4-carriers: A replication study.* NeuroImage, 2021. **240**: p. 118304.
12. Insel, P.S., O. Hansson, and N. Mattsson-Carlgren, *Association Between Apolipoprotein E ε2 vs ε4, Age, and β-Amyloid in Adults Without Cognitive Impairment.* JAMA Neurology, 2021. **78**(2): p. 229-235.
13. Braak, E. and H. Braak, *Alzheimer's disease: transiently developing dendritic changes in pyramidal cells of sector CA1 of the Ammon's horn.* Acta Neuropathologica, 1997. **93**(4): p. 323-325.
14. Shafer, A.T., et al., *Default mode network connectivity and cognition in the aging brain: the effects of age, sex, and APOE genotype.* Neurobiology of Aging, 2021. **104**: p. 10-23.
15. Salami, A., S. Pudas, and L. Nyberg, *Elevated hippocampal resting-state connectivity underlies deficient neurocognitive function in aging.* Proceedings of the National Academy of Sciences, 2014. **111**(49): p. 17654.
16. Kobayashi, Y. and D.G. Amaral, *Macaque monkey retrosplenial cortex: III. Cortical efferents.* Journal of Comparative Neurology, 2007. **502**(5): p. 810-833.
17. Aggleton, J.P., et al., *Complementary Patterns of Direct Amygdala and Hippocampal Projections to the Macaque Prefrontal Cortex.* Cerebral Cortex, 2015. **25**(11): p. 4351-4373.

18. Barbas, H. and G.J. Blatt, *Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey.* Hippocampus, 1995. **5**(6): p. 511-533.
19. Fleisher, A., et al., *Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment.* Neurology, 2008. **70**(3): p. 191-199.
20. Braak, H. and E. Braak, *Neuropathological stageing of Alzheimer-related changes.* Acta neuropathologica, 1991. **82**(4): p. 239-259.
21. Iglesias, J.E., 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**: p. 117-137.
22. Farrer, L.A., 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): p. 1349-1356.
23. Wang, H.-T., et al., *Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists.* NeuroImage, 2020. **216**: p. 116745.
24. Zajner, C., R.N. Spreng, and D. Bzdok, *Loneliness is linked to specific subregional alterations in hippocampus-default network covariation.* Journal of Neurophysiology, 2021. **126**(6): p. 2138-2157.
25. Padurariu, M., et al., *Hippocampal neuronal loss in the CA1 and CA3 areas of Alzheimer's disease patients.* Psychiatria Danubina, 2012. **24**(2.): p. 152-158.
26. Gold, B.T., et al., *Alterations in multiple measures of white matter integrity in normal women at high risk for Alzheimer's disease.* Neuroimage, 2010. **52**(4): p. 1487-1494.
27. Mielke, M.M., et al., *Fornix integrity and hippocampal volume predict memory decline and progression to Alzheimer's disease.* Alzheimer's & Dementia, 2012. **8**(2): p. 105-113.
28. Badhwar, A., et al., *Impaired structural correlates of memory in Alzheimer's disease mice.* NeuroImage: Clinical, 2013. **3**: p. 290-300.
29. Palmqvist, S., et al., *Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity.* Nature communications, 2017. **8**(1): p. 1-13.
30. Damoiseaux, J.S., et al., *Gender modulates the APOE ε4 effect in healthy older adults: convergent evidence from functional brain connectivity and spinal fluid tau levels.* Journal of Neuroscience, 2012. **32**(24): p. 8254-8262.
31. Reiman, E.M., et al., *Correlations between apolipoprotein E ε4 gene dose and brain-imaging measurements of regional hypometabolism.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(23): p. 8299-8302.
32. Reiman, E.M., et al., *Preclinical Evidence of Alzheimer's Disease in Persons Homozygous for the ε4 Allele for Apolipoprotein E.* New England Journal of Medicine, 1996. **334**(12): p. 752-758.
33. Corder, E., et al., *Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease.* Nature genetics, 1994. **7**(2): p. 180-184.
34. Shafighi, K., et al., *Social isolation is linked to classical risk factors of Alzheimer's disease-related dementias.* bioRxiv, 2021.
35. Nasreddine, Z.S., et al., *The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment.* Journal of the American Geriatrics Society, 2005. **53**(4): p. 695-699.

36. Li, M., et al., *Altered Hippocampal Subfields Volumes Is Associated With Memory Function in Type 2 Diabetes Mellitus.* Front Neurol, 2021. **12**: p. 756500.
37. Lim, Y.Y., E.C. Mormino, and A.s.D.N. Initiative, *APOE genotype and early β-amyloid accumulation in older adults without dementia.* Neurology, 2017. **89**(10): p. 1028-1034.
38. Nagy, Z.S., et al., *Influence of the apolipoprotein E genotype on amyloid deposition and neurofibrillary tangle formation in Alzheimer's disease.* Neuroscience, 1995. **69**(3): p. 757-761.
39. Reiman, E.M., et al., *Exceptionally low likelihood of Alzheimer's dementia in APOE2 homozygotes from a 5,000-person neuropathological study.* Nature Communications, 2020. **11**(1): p. 667.
40. Jansen, W.J., et al., *Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia: A Meta-analysis.* JAMA, 2015. **313**(19): p. 1924-1938.
41. Kim, Y.J., et al., *Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: a three-year longitudinal study.* Scientific reports, 2017. **7**(1): p. 1-8.
42. Lippa, C.F., et al., *Apolipoprotein E-epsilon 2 and Alzheimer's disease.* Genotype influences pathologic phenotype, 1997. **48**(2): p. 515-519.
43. Shu, H., et al., *Opposite Neural Trajectories of Apolipoprotein E ϵ4 and ϵ2 Alleles with Aging Associated with Different Risks of Alzheimer's Disease.* Cerebral Cortex, 2016. **26**(4): p. 1421-1429.
44. Shinohara, M., et al., *APOE2 eases cognitive decline during aging: clinical and preclinical evaluations.* Annals of neurology, 2016. **79**(5): p. 758-774.
45. Riddell, D.R., et al., *Impact of Apolipoprotein E (ApoE) Polymorphism on Brain ApoE Levels.* The Journal of Neuroscience, 2008. **28**(45): p. 11445-11453.
46. Ulrich, J.D., et al., *In vivo measurement of apolipoprotein E from the brain interstitial fluid using microdialysis.* Molecular Neurodegeneration, 2013. **8**(1): p. 13.
47. Beffert, U., 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 Research, 1999. **843**(1): p. 87-94.
48. Krasemann, S., et al., *The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases.* Immunity, 2017. **47**(3): p. 566-581.e9.
49. Huynh, T.-P.V., et al., *Lack of hepatic apoE does not influence early Aβ deposition: observations from a new APOE knock-in model.* Molecular neurodegeneration, 2019. **14**(1): p. 1-23.
50. Deane, R., et al., *apoE isoform–specific disruption of amyloid β peptide clearance from mouse brain.* The Journal of clinical investigation, 2008. **118**(12): p. 4002-4013.
51. Berlau, D.J., et al., *APOE ε2 is associated with intact cognition but increased Alzheimer pathology in the oldest old.* Neurology, 2009. **72**(9): p. 829-834.
52. Galea, L.A., et al., *Sex, hormones and neurogenesis in the hippocampus: hormonal modulation of neurogenesis and potential functional implications.* Journal of neuroendocrinology, 2013. **25**(11): p. 1039-1061.
53. Galea, L. and B. McEwen, *Sex and seasonal changes in the rate of cell proliferation in the dentate gyrus of adult wild meadow voles.* Neuroscience, 1999. **89**(3): p. 955-964.

54. Tanapat, P., et al., *Estrogen stimulates a transient increase in the number of new neurons in the dentate gyrus of the adult female rat.* Journal of Neuroscience, 1999. **19**(14): p. 5792-5801.
55. Stone, D.J., et al., *Astrocytes and Microglia Respond to Estrogen with Increased apoE mRNAin Vivoandin Vitro.* Experimental Neurology, 1997. **143**(2): p. 313-318.
56. Lambert, J., N. Coyle, and C. Lendon, *The allelic modulation of apolipoprotein E expression by oestrogen: potential relevance for Alzheimer's disease.* Journal of medical genetics, 2004. **41**(2): p. 104-112.
57. Mahley, R.W., Y. Huang, and S.C. Rall, *Pathogenesis of type III hyperlipoproteinemia (dysbetalipoproteinemia): questions, quandaries, and paradoxes.* Journal of lipid research, 1999. **40**(11): p. 1933-1949.
58. Metzler-Baddeley, C., et al., *Sex-specific effects of central adiposity and inflammatory markers on limbic microstructure.* NeuroImage, 2019. **189**: p. 793-803.
59. Sochocka, M., et al., *The Gut Microbiome Alterations and Inflammation-Driven Pathogenesis of Alzheimer's Disease—a Critical Review.* Molecular Neurobiology, 2019. **56**(3): p. 1841-1851.
60. Kernbach, J.M., et al., *Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants.* Proceedings of the National Academy of Sciences, 2018. **115**(48): p. 12295-12300.
61. Yamatani, H., et al., *Association of estrogen with glucocorticoid levels in visceral fat in postmenopausal women.* Menopause, 2013. **20**(4).
62. Nelson, P.T., et al., *APOE-ε2 and APOE-ε4 Correlate With Increased Amyloid Accumulation in Cerebral Vasculature.* Journal of Neuropathology & Experimental Neurology, 2013. **72**(7): p. 708-715.
63. Yu, L., et al., *APOE and cerebral amyloid angiopathy in community-dwelling older persons.* Neurobiology of Aging, 2015. **36**(11): p. 2946-2953.
64. McKay, G.J., et al., *Evidence of association of APOE with age-related macular degeneration-a pooled analysis of 15 studies.* Human mutation, 2011. **32**(12): p. 1407-1416.
65. Zhao, N., et al., *APOE ε2 is associated with increased tau pathology in primary tauopathy.* Nature communications, 2018. **9**(1): p. 1-11.
66. Ghebremedhin, E., et al., *Argyrophilic grain disease is associated with apolipoprotein E ε2 allele.* Acta neuropathologica, 1998. **96**(3): p. 222-224.
67. Mahley, R.W., *Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders.* Journal of Molecular Medicine, 2016. **94**(7): p. 739-746.
68. Wang, H.-X., et al., *Late-life engagement in social and leisure activities is associated with a decreased risk of dementia: a longitudinal study from the Kungsholmen project.* American journal of epidemiology, 2002. **155**(12): p. 1081-1087.
69. Podewils, L.J., et al., *Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study.* American journal of epidemiology, 2005. **161**(7): p. 639-651.
70. Rovio, S., et al., *Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease.* The Lancet Neurology, 2005. **4**(11): p. 705-711.
71. Grande, G., et al., *Physical activity reduces the risk of dementia in mild cognitive impairment subjects: a cohort study.* Journal of Alzheimer's disease, 2014. **39**(4): p. 833-839.

72. Tan, Z.S., et al., *Physical activity, brain volume, and dementia risk: the Framingham study.* The Journals of Gerontology: Series A, 2017. **72**(6): p. 789-795.
73. Schächter, F., et al., *Genetic associations with human longevity at the APOE and ACE loci.* Nature Genetics, 1994. **6**(1): p. 29-32.
74. Ranft, U., et al., *Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly.* Environmental Research, 2009. **109**(8): p. 1004-1011.
75. Wellenius, G.A., et al., *Residential Proximity to Nearest Major Roadway and Cognitive Function in Community-Dwelling Seniors: Results from the MOBILIZE Boston Study.* Journal of the American Geriatrics Society, 2012. **60**(11): p. 2075-2080.
76. Weuve, J., et al., *Exposure to Particulate Air Pollution and Cognitive Decline in Older Women.* Archives of Internal Medicine, 2012. **172**(3): p. 219-227.
77. Ailshire, J.A. and E.M. Crimmins, *Fine Particulate Matter Air Pollution and Cognitive Function Among Older US Adults.* American Journal of Epidemiology, 2014. **180**(4): p. 359-366.
78. Ailshire, J.A. and P. Clarke, *Fine Particulate Matter Air Pollution and Cognitive Function Among U.S. Older Adults.* The Journals of Gerontology: Series B, 2014. **70**(2): p. 322-328.
79. Tonne, C., et al., *Traffic-related air pollution in relation to cognitive function in older adults.* Epidemiology (Cambridge, Mass.), 2014. **25**(5): p. 674-681.
80. Gatto, N.M., et al., *Components of air pollution and cognitive function in middle-aged and older adults in Los Angeles.* NeuroToxicology, 2014. **40**: p. 1-7.
81. Tzivian, L., 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.* Environmental Health Perspectives, 2016. **124**(9): p. 1361-1368.
82. Cacciottolo, M., 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): p. e1022.
83. Jung, C.-R., Y.-T. Lin, and B.-F. Hwang, *Ozone, Particulate Matter, and Newly Diagnosed Alzheimer's Disease: A Population-Based Cohort Study in Taiwan.* Journal of Alzheimer's Disease, 2015. **44**: p. 573-584.
84. Wu, Y.-C., et al., *Association between air pollutants and dementia risk in the elderly.* Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 2015. **1**(2): p. 220-228.
85. Schikowski, T., et al., *Association of air pollution with cognitive functions and its modification by APOE gene variants in elderly women.* Environmental Research, 2015. **142**: p. 10-16.
86. Oudin, A., et al., *Traffic-Related Air Pollution and Dementia Incidence in Northern Sweden: A Longitudinal Study.* Environmental Health Perspectives, 2016. **124**(3): p. 306-312.
87. Chen, H., et al., *Living near major roads and the incidence of dementia, Parkinson's disease, and multiple sclerosis: a population-based cohort study.* The Lancet, 2017. **389**(10070): p. 718-726.
88. Zanobetti, A., et al., *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.* Environmental Health, 2014. **13**(1): p. 1-11.

89. Kioumourtzoglou, M.-A., et al., *Long-term PM2. 5 exposure and neurological hospital admissions in the northeastern United States.* Environmental health perspectives, 2016. **124**(1): p. 23-29.
90. Carey, I.M., et al., *Are noise and air pollution related to the incidence of dementia? A cohort study in London, England.* BMJ open, 2018. **8**(9): p. e022404.
91. Chang, K.-H., et al., *Increased risk of dementia in patients exposed to nitrogen dioxide and carbon monoxide: a population-based retrospective cohort study.* PloS one, 2014. **9**(8): p. e103078.
92. Chen, H., et al., *Exposure to ambient air pollution and the incidence of dementia: a population-based cohort study.* Environment international, 2017. **108**: p. 271-277.
93. Calderón-Garcidueñas, L., 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.* Environmental research, 2016. **146**: p. 404-417.
94. Sengillo, J.D., et al., *Deficiency in Mural Vascular Cells Coincides with Blood–Brain Barrier Disruption in A lzheimer's Disease.* Brain pathology, 2013. **23**(3): p. 303-310.
95. Koutsilieri, E., et al., *Degeneration of neuronal cells due to oxidative stress—microglial contribution.* Parkinsonism & Related Disorders, 2002. **8**(6): p. 401-406.
96. Roqué, P.J., K. Dao, and L.G. Costa, *Microglia mediate diesel exhaust particle-induced cerebellar neuronal toxicity through neuroinflammatory mechanisms.* NeuroToxicology, 2016. **56**: p. 204-214.
97. Giordano, G., et al., *Gender differences in brain susceptibility to oxidative stress are mediated by levels of paraoxonase-2 expression.* Free Radical Biology and Medicine, 2013. **58**: p. 98-108.
98. Costa, L.G., et al., *Paraoxonase-2 (PON2) in brain and its potential role in neuroprotection.* NeuroToxicology, 2014. **43**: p. 3-9.
99. Borrás, C., et al., *Mitochondria from females exhibit higher antioxidant gene expression and lower oxidative damage than males.* Free Radical Biology and Medicine, 2003. **34**(5): p. 546-552.
100. Raichle, M.E., et al., *A default mode of brain function.* Proceedings of the National Academy of Sciences, 2001. **98**(2): p. 676-682.
101. Liguori, C., et al., *Cerebrospinal fluid lactate levels and brain [18F] FDG PET hypometabolism within the default mode network in Alzheimer's disease.* European journal of nuclear medicine and molecular imaging, 2016. **43**(11): p. 2040-2049.
102. Malkov, A., et al., *Aβ initiates brain hypometabolism, network dysfunction and behavioral abnormalities via NOX2-induced oxidative stress in mice.* Communications Biology, 2021. **4**(1): p. 1-12.
103. Daftary, S.S. and A.C. Gore, *IGF-1 in the brain as a regulator of reproductive neuroendocrine function.* Experimental biology and medicine, 2005. **230**(5): p. 292-306.
104. Åberg, M.A.I., et al., *Peripheral Infusion of IGF-I Selectively Induces Neurogenesis in the Adult Rat Hippocampus.* The Journal of Neuroscience, 2000. **20**(8): p. 2896-2903.
105. Trejo, J.L., E. Carro, and I. Torres-Alemán, *Circulating Insulin-Like Growth Factor I Mediates Exercise-Induced Increases in the Number of New Neurons in the Adult Hippocampus.* The Journal of Neuroscience, 2001. **21**(5): p. 1628-1634.

106. Altman, J. and G.D. Das, *Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats.* Journal of Comparative Neurology, 1965. **124**(3): p. 319-335.
107. Kaplan, M.S. and J.W. Hinds, *Neurogenesis in the adult rat: electron microscopic analysis of light radioautographs.* Science, 1977. **197**(4308): p. 1092-1094.
108. Kornack, D.R. and P. Rakic, *Continuation of neurogenesis in the hippocampus of the adult macaque monkey.* Proceedings of the National Academy of Sciences, 1999. **96**(10): p. 5768-5773.
109. Hagenaars, S.P., et al., *Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia.* Molecular Psychiatry, 2016. **21**(11): p. 1624-1632.
110. Lyall, D.M., et al., *Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants.* Age and ageing, 2016. **45**(4): p. 511-517.
111. Davies, G., et al., *Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N= 112 151).* Molecular psychiatry, 2016. **21**(6): p. 758-767.
112. Tank, R., et al., *Association between polygenic risk for Alzheimer's disease, brain structure and cognitive abilities in UK Biobank.* Neuropsychopharmacology, 2022. **47**(2): p. 564-569.
113. Bartzokis, G., *Alzheimer's disease as homeostatic responses to age-related myelin breakdown.* Neurobiology of Aging, 2011. **32**(8): p. 1341-1371.
114. Tavares, R.M., et al., *A map for social navigation in the human brain.* Neuron, 2015. **87**(1): p. 231-243.
115. Morrish, N. and A. Medina-Lara, *Does unemployment lead to greater levels of loneliness? A systematic review.* Social Science & Medicine, 2021: p. 114339.
116. Penninkilampi, R., et al., *The Association between Social Engagement, Loneliness, and Risk of Dementia: A Systematic Review and Meta-Analysis.* Journal of Alzheimer's Disease, 2018. **66**: p. 1619-1633.
117. Kuiper, J.S., et al., *Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies.* Ageing research reviews, 2015. **22**: p. 39-57.
118. Bassuk, S.S., T.A. Glass, and L.F. Berkman, *Social disengagement and incident cognitive decline in community-dwelling elderly persons.* Annals of internal medicine, 1999. **131**(3): p. 165-173.
119. Scarmeas, N., et al., *Influence of leisure activity on the incidence of Alzheimer's disease.* Neurology, 2001. **57**(12): p. 2236-2242.
120. Hostinar, C.E. and M.R. Gunnar, *Social support can buffer against stress and shape brain activity.* AJOB neuroscience, 2015. **6**(3): p. 34-42.
121. Cacioppo, J.T., et al., *The Neuroendocrinology of Social Isolation.* Annual Review of Psychology, 2015. **66**(1): p. 733-767.
122. Bzdok, D. and R.I. Dunbar, *The neurobiology of social distance.* Trends in cognitive sciences, 2020. **24**(9): p. 717-733.
123. Kong, X., et al., *Neuroticism and extraversion mediate the association between loneliness and the dorsolateral prefrontal cortex.* Experimental brain research, 2015. **233**(1): p. 157-164.
124. Lecic-Tosevski, D., O. Vukovic, and J. Stepanovic, *Stress and personality.* Psychiatriki, 2011. **22**(4): p. 290-297.

125. Swanson, L.W. and W.M. Cowan, *Hippocampo-Hypothalamic Connections: Origin in Subicular Cortex, Not Ammon's Horn.* Science, 1975. **189**(4199): p. 303-304.
126. Wilson, R.S., et al., *Chronic distress and incidence of mild cognitive impairment.* Neurology, 2007. **68**(24): p. 2085-2092.
127. Wilson, R., et al., *Proneness to psychological distress is associated with risk of Alzheimer's disease.* Neurology, 2003. **61**(11): p. 1479-1485.
128. Wilson, R., et al., *Proneness to psychological distress and risk of Alzheimer disease in a biracial community.* Neurology, 2005. **64**(2): p. 380-382.
129. Wilson, R.S., et al., *Chronic psychological distress and risk of Alzheimer's disease in old age.* Neuroepidemiology, 2006. **27**(3): p. 143-153.
130. Fiocco, A., 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.* Neurobiology of aging, 2009. **30**(6): p. 1001-1005.
131. Levy, B.R., et al., *When Culture Influences Genes: Positive Age Beliefs Amplify the Cognitive-Aging Benefit of APOE ε2.* The Journals of Gerontology: Series B, 2020. **75**(8): p. e198-e203.
132. Alfaro-Almagro, F., et al., *Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank.* Neuroimage, 2018. **166**: p. 400-424.
133. Zhang, Q., et al., *Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture.* Nature communications, 2020. **11**(1): p. 1-11.
134. Liu, J.Z., Y. Erlich, and J.K. Pickrell, *Case–control association mapping by proxy using family history of disease.* Nature Genetics, 2017. **49**(3): p. 325-331.
135. Marioni, R.E., et al., *GWAS on family history of Alzheimer's disease.* Translational psychiatry, 2018. **8**(1): p. 1-7.
136. Kunkle, B.W., et al., *Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing.* Nature Genetics, 2019. **51**(3): p. 414-430.
137. Lambert, J.-C., et al., *Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease.* Nature Genetics, 2013. **45**(12): p. 1452-1458.
138. Miller, K.L., et al., *Multimodal population brain imaging in the UK Biobank prospective epidemiological study.* Nature neuroscience, 2016. **19**(11): p. 1523.
139. Smith, S.M., *Fast robust automated brain extraction.* Hum Brain Mapp, 2002. **17**(3): p. 143-55.
140. Jenkinson, M. and S. Smith, *A global optimisation method for robust affine registration of brain images.* Medical Image Analysis, 2001. **5**(2): p. 143-156.
141. Jenkinson, M., et al., *Improved optimization for the robust and accurate linear registration and motion correction of brain images.* Neuroimage, 2002. **17**(2): p. 825-841.
142. Andersson, J.L., M. Jenkinson, and S. Smith, *Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2.* FMRIB Analysis Group of the University of Oxford, 2007. **2**(1): p. e21.
143. Zhang, Y., M. Brady, and S. Smith, *Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm.* IEEE Transactions on Medical Imaging, 2001. **20**(1): p. 45-57.
144. Smith, S.M., et al., *Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis.* NeuroImage, 2002. **17**(1): p. 479-489.

145. Schaefer, A., et al., *Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI.* Cerebral Cortex, 2017: p. 1-20.
146. Spreng, R.N., et al., *The default network of the human brain is associated with perceived social isolation.* Nature communications, 2020. **11**(1): p. 1-11.
147. Schurz, M., et al., *Variability in brain structure and function reflects lack of peer support.* Cerebral Cortex, 2021.
148. Witten, D.M., R. Tibshirani, and T. Hastie, *A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis.* Biostatistics, 2009: p. kxp008.
149. Bzdok, D., T.E. Nichols, and S.M. Smith, *Towards algorithmic analytics for large-scale datasets.* Nature Machine Intelligence, 2019. **1**(7): p. 296-306.
150. Zajner, C., R.N. Spreng, and D. Bzdok, *Lacking social support is associated with structural divergences in hippocampus–default network co-variation patterns.* Social Cognitive and Affective Neuroscience, 2022. **17**(9): p. 802-818.
151. Efron, B. and R.J. Tibshirani, *An introduction to the bootstrap*. 1994: CRC press.
152. Millard, L.A.C., et al., *Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank.* Int J Epidemiol, 2018. **47**(1): p. 29-35.
153. Wolpert, D.H., *Stacked generalization.* Neural networks, 1992. **5**(2): p. 241-259.
154. Breiman, L., *Stacked regressions.* Machine learning, 1996. **24**(1): p. 49-64.
155. Dong, Q., et al., *Applying surface-based hippocampal morphometry to study APOE-E4 allele dose effects in cognitively unimpaired subjects.* NeuroImage: Clinical, 2019. **22**: p. 101744.
156. Hostage, C.A., et al., *Dissecting the gene dose-effects of the APOE ε4 and ε2 alleles on hippocampal volumes in aging and Alzheimer's disease.* PloS one, 2013. **8**(2): p. e54483.
157. West, H.L., G.W. Rebeck, and B.T. Hyman, *Frequency of the apolipoprotein E ε2 allele is diminished in sporadic Alzheimer disease.* Neuroscience letters, 1994. **175**(1-2): p. 46-48.
158. Benjamin, R., et al., *Protective effect of apoE epsilon 2 in Alzheimer's disease.* Lancet (London, England), 1994. **344**(8920): p. 473-473.
159. Lecarpentier, J., et al., *Prediction of Breast and Prostate Cancer Risks in Male BRCA1 and BRCA2 Mutation Carriers Using Polygenic Risk Scores.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2017. **35**(20): p. 2240-2250.
160. Fan, B.J., et al., *Association of a Primary Open-Angle Glaucoma Genetic Risk Score With Earlier Age at Diagnosis.* JAMA Ophthalmology, 2019. **137**(10): p. 1190-1194.
161. Benjamini, Y. and Y. Hochberg, *Controlling the false discovery rate: a practical and powerful approach to multiple testing.* Journal of the Royal statistical society: series B (Methodological), 1995. **57**(1): p. 289-300.
162. Sha, Z., et al., *The genetic architecture of structural left-right asymmetry of the human brain.* Nat Hum Behav, 2021. **5**(9): p. 1226-1239.
163. Raizada, R.D., et al., *Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children.* Neuroimage, 2008. **40**(3): p. 1392-401.
164. Genovese, C.R., N.A. Lazar, and T. Nichols, *Thresholding of statistical maps in functional neuroimaging using the false discovery rate.* Neuroimage, 2002. **15**(4): p. 870-8.

165. Ward Jr, J.H., *Hierarchical grouping to optimize an objective function.* Journal of the American statistical association, 1963. **58**(301): p. 236-244.
166. Pearson, K., *LIII. On lines and planes of closest fit to systems of points in space.* The London, Edinburgh, and Dublin philosophical magazine and journal of science, 1901. **2**(11): p. 559-572.
167. Gerbrands, J.J., *On the relationships between SVD, KLT and PCA.* Pattern recognition, 1981. **14**(1-6): p. 375-381.
168. Tremblay-Mercier, J., et al., *Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease.* NeuroImage: Clinical, 2021. **31**: p. 102733.

Bridge

By harnessing the capabilities of machine learning and big data analytics, our objective is to furnish an impartial, data-driven portrayal of the phenotypic divergence in familial AD risk, considering sex differences, across two substantial cohorts of asymptomatic older adults. In the initial phase detailed in Chapter 1, we leveraged the comprehensive and uniformly collected brain imaging data from the UK Biobank (UKB) to derive population insights into neuroanatomical correlates of Alzheimer's disease and related dementias (ADRD) family risk. This endeavour sought to establish sex-specific phenome-wide profiles of ADRD vulnerability associated with micro-anatomically defined subregions within the hippocampus (HC) and default network (DN)—two brain systems recognized for exhibiting early aberrations along the AD continuum [12, 44, 69, 70]. Employing a curated set of approximately 1,000 phenotypes, we furnished a comprehensive overview of the incipient manifestations of ADRD in relation to sex, *APOE* polymorphism, and the co-variation between the allocortical HC and neocortical DN.

Working with the UKB made it uniquely possible for us to derive distinct patterns of structural co-variation between 38 subregions of the HC and 91 subregions of the DN, which is unachievable with the overwhelming majority of other brain-imaging samples. A recent comprehensive examination of the performance of canonical correlation analysis (CCA)—the core model of our present neuroanatomical investigation—in biomedical datasets has shown that perhaps only the UKB offers large enough participant sample sizes to obtain stable estimates of our doubly-multivariate model (i.e., CCA) in the high-dimensional data setting [72]. Brain-imaging datasets typically considered "large" according to community standards, such as the Human Connectome Project ($n \approx 1000$), were prone to overfitting and did not reach convergence to stable parameter estimates for a set of 100 features. When the number of subjects is too close to the number of features, CCA struggles to approximate any valid latent dimensions—it fails to find a unique identifiable modelling solution [73]. While the UKB was the only brain imaging resource with the power to enable such a highly multivariate decomposition, it was not developed with the aim of tracking ADRD progression. For that reason, the biological and phenotypic markers investigated by the UKB, although covering a wide range of health-related determinants, are rather unspecific to dementia progression.

In Chapter 2, we will aim to build upon the characterization of our population-derived limbic-cortical regimes by linking them with widely established indicators of dementia progression in asymptomatic participants with Alzheimer's disease (AD) family history. Precisely, we will carry over and investigate the expressions of our HC-DN signatures in a specialized cohort of individuals at heightened risk of developing dementia that offers one of the most rigorous monitoring of AD progression in offspring of diagnosed patients: the Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD (PREVENT-AD) cohort (Tremblay-Mercier, Madjar et al. 2021). This rich prospective cohort contains annual assessments of most, if not all, major AD markers, including blood and CSF biochemistry, cardiovascular, neurosensory, and cognitive indicators, as well as medical and demographic records. By utilizing these two distinct datasets, we aim to enhance our comprehension of the comparability between broadly applicable facets of participants' health and clinically significant indicators of AD risk by how robustly they can be tracked by HC-DN co-variation. Beyond the external validation of our outcomes, this secondary cohort facilitates a focused exploration of intergenerational variances in AD heritability with respect to sex. Our precise endeavour is to investigate how the sex of both the at-risk progeny and the parent diagnosed with AD influences the preclinical disease manifestations across established and clinically relevant biological markers of AD progression. Given the deep phenotyping offered by the PREVENT-AD resource, our investigation was uniquely positioned to chart maternal vs. paternal risk effects across the whole at-risk phenome—something out of reach in most clinical datasets. Our novel approach builds on population-level insights into structural deviation patterns in the DN and HC to now quantify the extent to which maternal and paternal lineage is reflected in AD-vulnerable brain structures. This two-step pipeline, empowered by the robustness of the UKB and the precision of PREVENT-AD, is positioned to reveal unprecedented insights into generational sex biases in AD symptomatology and its neuroanatomical underpinnings.

Chapter 2: Parent-of-origin effects in Alzheimer's liability dissociate neurocognitive and cardiovascular traits in at-risk individuals

Chloé Savignac[1,2], Frédéric St-Onge[2,3], Sylvia Villeneuve[3,4,5,6,7], AmanPreet Badhwar[8,9], Sarah A. Gagliano Taliun[10,11], Sali Farhan[5,12], Maiya Geddes[3,4,5,13,14], Yasser Iturria Medina[5,6,15], Judes Poirier[3,4,5,7], R. Nathan Spreng[3,4,5,6,7,16], Danilo Bzdok[1,6,17,18], the PREVENT-AD Research Group*

1 Department of Biomedical Engineering, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

2 Integrated Program in Neuroscience, Faculty of medicine, McGill University, Montreal, Quebec, Canada

3 Research center of the Douglas Mental Health Institute, Montreal, Quebec, Canada

4 Centre for Studies in the Prevention of Alzheimer's Disease, Douglas Mental Health Institute, McGill University, Montreal, Quebec, Canada

5 Department of Neurology and Neurosurgery, Montreal Neurological Institute (MNI), Faculty of Medicine, McGill University, Montreal, Quebec, Canada.

6 McConnell Brain Imaging Centre (BIC), MNI, Faculty of Medicine, McGill University, Montreal, Quebec, Canada.

7 Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

8 Département de pharmacologie et physiologie & Institut de génie biomédical, Faculté de médecine, Université de Montréal, Montreal, Quebec, Canada

9 Centre de recherche de l'Institut universitaire de gériatrie de Montréal (CRIUGM), Montreal, Quebec, Canada

10 Department of neurosciences & Department of medicine, Faculty of medicine, Université de Montréal, Montreal, Quebec, Canada

11 Montreal Heart Institute, Montréal, Quebec, Canada

12 Department of Human Genetics, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

13 Research Centre for Studies in Aging, Douglas Mental Health Institute, McGill University, Montreal, Quebec, Canada

14 McGill University Research Centre for Studies in Aging, Montreal, Quebec, Canada

15 Ludmer Centre for Neuroinformatics and Mental Health, McGill University, Montreal, Quebec, Canada

16 Department of Psychology, Faculty of Medicine, McGill University, Montreal, Quebec, Canada

17 School of Computer Science, McGill University, Montreal, Quebec, Canada

18 Mila – Quebec Artificial Intelligence Institute, Montreal, Quebec, Canada

* Data used in preparation of this article were obtained from the PRe-symptomatic EValuation of Experimental or Novel Treatments for Alzheimer's Disease (PREVENT-AD) program (https://douglas.research.mcgill.ca/stop-ad-centre), data release 6.0. A complete listing of PREVENT-AD Research Group can be found in the PREVENT-AD database: https://preventad.loris.ca/acknowledgements/acknowledgements.php?date=[YYYY-MM-DD] †.

Abstract:

Alzheimer's disease (AD) has a higher prevalence among women. Also, more women rather than men among AD patients pass on the disease to their children. Yet, evidence of father-to-son transmission has been documented by epidemiological studies, which challenges the dominant matrilinear narrative. Here, by means of phenome-wide assays, we aimed to reconcile clinical reports with population-based insights as to diverging influences of maternal vs. paternal AD risk on first-degree relatives. We capitalized on the richest single-site cohort (PREVENT-AD) with a family history of AD to extract three distinct intermediate phenotypes of AD susceptibility across the PREVENT-AD phenome. Capitalizing on ~1,000 individual subject visits, we examined how much the derived intermediate phenotypes vary as a function on maternal vs. paternal AD lineage. Our careful cross-generational examination highlighted the influences of matri- vs. patrilinear AD risk on cardiovascular and cognitive risk. Notably, we identified sex bias for polymorphisms in the *HMGCR* and *BDNF* genes as most explanatory for AD genealogy passing through the mother vs. father. Zooming in on microanatomical alterations in hippocampus and default network subregions, we identified distinct structural patterns related to matri- vs. patrilinear AD risk in subregions from which the fornix white-matter tract originates. As the most systematic study of its kind, our cross-generational analysis ultimately delineated parents-of-origin effects in AD genealogy.

Keywords: lineage, phenomics, PREVENT-AD, Alzheimer's disease, family risk, parents-of-origin effects

Introduction

Parental history of sporadic Alzheimer's disease (AD) has been recognized as a key disease risk factor since the late 1980s [1]. The cumulative risk of developing dementia is ~20-65% higher for the offspring of AD patients than for individuals without a family history [2]. Sex, in turn, has been considered a non-negligible risk factor for developing AD-type dementia for at least two decades [2]. AD prevalence is higher in female subjects [3]. Transmission of AD risk is also more frequent in offspring with maternal rather than paternal history of AD [3]. Important for our present investigation, maternal inheritance is thought to be 1.7 to 3.6 times more frequent than paternal inheritance [4].

Despite these well-established sex divergences in AD genealogy, an early brain imaging study only recently compared maternal vs. paternal AD risk in 11 males and 13 females with a family history [5]. Since then, the brain imaging community has sporadically studied small clinical samples of at-risk individuals (typically 8-16 subjects [5-8]) to investigate the neurobiological correlates of AD familial risk. Importantly, working with such handpicked clinical samples did not allow researchers to peel apart the effects of maternal vs. paternal risk on male and female offspring. The statistical power, arguably too modest, often pushed for case-control investigations in which neither the sexes of the AD patients nor at-risk children were considered. As a result, the potentially diverging clinical manifestations of maternal and paternal AD risk in at-risk offspring remain largely unstudied. This is despite longstanding epidemiological evidence for sex-specific transmission of AD liability.

Due to lower prevalence and narrow clinical samples, there has been limited appreciation of a paternal family history of AD – yet it may turn out to be a potentially separate risk phenotype. Indeed, recent epidemiological population estimates from more than a million participants – inaccessible to brain-imaging assessments at that scale – suggested that men are more vulnerable to developing dementia than women with the same family history [9]. Somewhat counterintuitively, paternal family history of AD has widely been regarded as a control group, similar to subjects with negative parental history, rather than studied as a disease risk category in its own right. Against a background of paternal risk, previous studies have established the effects

of maternal inheritance on brain volume [10], mitochondrial activity [11], glucose metabolism [7], and cerebral blood flow [12]. This common practice even reached a point where paternal and negative family history of AD were combined in data visualization [7] and classification analyses [13] to spotlight maternal effects. While population-based estimates have highlighted male-specific generational risk, most small to medium size brain imaging experiments in a handful of participants have placed an overwhelming focus on the maternal transmission of AD risk.

By designing a data-driven framework, we conducted the, so far, most systematic head-to-head comparison of matrilinear (i.e., mother had AD) vs. patrilinear (i.e., father had AD) effects in the offspring of AD subjects, before the onset of any clinical symptoms. We capitalized on the largest-of-its-kind at-risk AD cohort, with ~1,000 visits of individuals from a family with a family history of AD: Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD (PREVENT-AD [14]). PREVENT-AD tracks AD onset and trajectory in cognitively unimpaired first-degree relatives of dementia patients. We built two sex-specific partial least square regression models in which the *APOE* genotypes (e.g., ε3/2, ε3/3, ε3/4) were estimated based on a collection of 256 PREVENT-AD risk indicators. This susceptible population has a 2-3-fold relative increase in the risk of AD-type dementia [15, 16]. Annual screening of the healthy PREVENT-AD participants was scheduled to monitor AD-specific fluctuations on well-established multimodal markers. The broad categories of risk factors offered by the deep phenotyping of PREVENT-AD allowed us to design a phenome-wide screening across a rich portfolio of 256 risk indicators: profiling of the probands' blood and cerebrospinal fluid (CSF) biochemistry, cardiovascular, neurosensory, and cognitive assessments, medical and demographics records, and brain scanning. The PREVENT-AD resource thus opened a unique window into how patrilinear vs. matrilinear AD inheritance differentially impacts risk phenotypes. We built on previous population-level insights into structural deviation patterns in the default network (DN) and hippocampus (HC) of healthy subjects with a family history of AD to quantify the extent to which maternal and paternal lineage was reflected in AD-vulnerable brain structures. Our novel approach to the inter-generational pathogenesis is thus positioned to reveal unprecedented insight into the parent-of-origin-specific nature of AD symptomatology.

Results

Rationale

Maternal history of AD is already a well-established dementia risk factor. AD has also been long known to entail partly diverging manifestations in males and females. It stands to reason that sex-specific disease processes could exercise a differential impact on the AD phenotype in the next generation. Nonetheless, matri- vs. patrilinear transmission of AD risk has not been systematically compared in male and female offspring. As a potential roadblock to progress, small handpicked clinical samples of typically ~20 subjects are limited in their appreciation of a paternal family history of AD, thus hindering assessment of a potentially distinct mode of inter-generational risk propagation. In parallel, population-based surveys of millions of participants are not equipped to perform a thorough clinical examination of well-established AD markers (e.g., CSF and blood biochemistry, grey matter volume, cognitive decline). To overcome several of these impasses, we leveraged the largest-of-its-kind family-based at-risk AD cohort to contrast matri- vs. patrilinear effects on AD risk in ~1,000 participant visits (PREVENT-AD). This rich prospective cohort contains annual assessment (mean= 65 years, ranging from 55 to 87) of most if not all major AD markers, including blood and CSF biochemistry, cardiovascular, neurosensory, and cognitive indicators, as well as medical and demographics records. Given this rich dataset, our investigation was uniquely positioned to chart matri- vs. patrilinear AD risk effects across the PREVENT-AD phenome. As a complementary analysis, we benefitted from image-derived phenotypes of grey matter morphology to assess whether maternal vs. paternal AD lineage was reflected in AD-vulnerable brain structures. We adopted a data-informed framework especially tailored for the systematic apples-to-apples comparison of AD risk in male vs. female subjects with maternal vs. paternal family history of AD — something out of reach in traditional AD clinical samples.

Intermediate phenotypes of AD susceptibility capture cis- and trans-generational variation in cognitive abilities and cardiovascular health

As the backbone of our analysis workflow, we extracted *APOE*-genotype-related intermediate phenotypes of AD susceptibility that captured different facets of AD familial risk by

being distinctly associated with ε3/2, ε3/3, and ε3/4 genotypes. We directly interrogated 256 PREVENT-AD risk indicators that spanned seven broad categories of AD risk factors: cardiovascular health, cognition, clinical co-morbidities, demographics, disease progression, genetics, and neurosensory assessments. Each derived intermediate phenotype thus encapsulated a different relationship between the three most prevalent *APOE* genotypes and the rich collection of PREVENT-AD risk indicators. Our next goal was to partition the variance in *APOE*-related AD risk effects across the PREVENT-AD phenome as a function of sex and maternal vs. paternal AD lineage. To do so, we examined how the distinct combinations of PREVENT-AD risk indicators captured by the three dominant intermediate phenotypes were differently expressed in males and females, as well as in individuals with maternal vs. paternal AD risk. We could thus quantify how much of the phenome-wide similarities in AD risk attributable to the *APOE* gene vary as a function of sex and maternal vs. paternal AD familial risk.

Our leading intermediate phenotype of AD susceptibility highlighted sex differences in genetic markers of memory performance and lipid metabolism. We found that the single nucleotide polymorphism (SNP) rs6265, also known as Val66Met (Fig. 1, right genetic radar plot), showed the most substantial sex bias. Val66Met is a missense variant in the gene that codes for the brain-derived neurotrophic factor (BDNF) protein. BDNF is a neurotrophic factor involved in synaptic plasticity and cognition [17]. The BDNF protein has been associated with several neurological disorders including AD due to its pivotal role in the integrity of hippocampal and neocortical neurons [18]. Post-mortem autopsy of AD patients has revealed decreased expression of BDNF mRNA in the hippocampus as compared to healthy controls [19]. The *BDNF* Val66Met polymorphism has been one of the most studied genetic variants in neurocognitive brain diseases over the past two decades [20]. Carriers of the Met allele have shown deficits in delayed and immediate recall as well as abnormal hippocampus activation [21]. The Met allele has also been associated with an increased risk of developing AD in females, but not in males [22, 23]. We located ValVal homozygotes and ValMet heterozygotes at opposite poles of the sex spectrum of disease variation for the first intermediate phenotype of AD susceptibility (Fig. 1) — consistent with the Val66Met polymorphism's reported sex bias. In addition to replicating sex-specific findings, we unveiled cross-generational differences in Val66Met polymorphism by highlighting a contingent influence of maternal vs. paternal AD liability. We have thus shown the effects of

BDNF polymorphism, a key marker of synaptic plasticity, to be driven by the sex of the AD at-risk subject, and secondarily also driven by the sex of the parent affected by AD.

The first phenotype of AD susceptibility also pointed to another genetic variant related to cardiovascular health. We found that SNP rs3846662, located in intron M of the 3-Hydroxy-3-Methylglutaryl-CoA Reductase (*HMGCR*) gene, was strongly dependent on family lineage. HMGCR is a key enzyme regulating cholesterol synthesis in mammalian cells in general [24]. *HMGCR* and *APOE* are thought to be involved in two separate yet interrelated pathways by which cholesterol is synthesized in the human brain [25, 26]. HMGCR is likely involved in the etiopathology of AD by regulating intracellular sterol sensing in the endoplasmic reticulum [27]. Being homozygous for the A allele (AA) in intron M of the *HMGCR* gene has been identified as one of the most important and common protective variants for sporadic AD, second only to *APOE* ε2 [28]. The A allele in intron M of the *HMGCR* gene possibly attenuates *APOE* ε4-mediated accumulation of hippocampal and neocortical amyloid plaques and tangles by acting on cholesterol synthesis in mammalian brain cells [28]. Amongst ε4 carriers, having two copies of the A allele has been shown to reduce AD conversion rate to a level similar to ε4 non-carriers [28]. Sex bias in *HMGCR* polymorphism has also been reported regarding the efficacy of statins [29]. A genotype-by-sex interaction was found such that women bearing the AA genotype showed increased overall transcription of *HMGCR* mRNA [29]. Similarly, post-mortem autopsies of AD patients have revealed increased *HMGCR* mRNA in the frontal cortex of women carrying the AA genotype [30]. Interrogating sex-specific variation in AD liability allowed us to add an additional level of complexity to the tight coupling between *APOE* and *HMGCR.* In particular, we located AA homozygotes and GA heterozygotes at opposite poles of the lineage spectrum of disease variation (Fig. 1, left genetic radar plot). *HMGCR* AA, but not GA, also showed a considerable variation in the sex spectrum of disease variation (Fig. 1, right genetic radar plot). Our cross-generational analysis captured the specific interplay between *APOE* ε4 and *HMGCR* A in a totally data-driven fashion. In fact, we found that genotype ε4/ε3 was driving most of the *APOE-HMGCR* associations as compared to genotypes ε3/3 and ε3/2 (Fig. 1). As such, pooling inter-generational variation in AD liability across *APOE* genotypes allowed us to single out *HMGCR* AA as driving sex- and lineage-specific variation in AD risk. Sex bias in *HMGCR* polymorphism has seldom been investigated with regard to AD liability. We have thus added to the previous *APOE-HMGCR*

studies by highlighting underlying generation and sex effects that were measurable in the AD at-risk subjects examined here.

The second intermediate phenotype of AD susceptibility unveiled notable effects of sex and AD lineage on cognitive markers, whereas neurosensory and cardiovascular indicators were predominantly influenced by sex. Age and cognitive indicators were situated at opposing extremes of the spectrum of disease variation with regard to both sex and AD lineage (Fig. 2; top left and bottom right corners). An additional distinction between markers of disordered vs. healthy olfaction was most evident with regard to sex. Total anosmia was situated near age indicators in the lower right corner, whereas normosmia and smell identification scores were located at the opposite end (Fig. 2; top left and bottom right corners). Indicators of cardiovascular health (e.g., systolic blood pressure) were positioned adjacent to age markers and exhibited influences from both sex and AD lineage. While olfactory loss can occur with certain neurogenerative diseases including AD, it can also be naturally impaired by aging [31]. Similarly, aging is a critical risk factor for all cardiovascular diseases known to substantially differs as a function of sex [32, 33]. With regards to *APOE* genetic risk, the second intermediate phenotype distinguished ε3 homozygotes from ε2 and ε4 carriers (Fig. 2). Commonly considered the baseline risk for AD, ε3 homozygotes could be seen as indicative of typical aging. However, we found a substantial influence of AD lineage on cognitive indicators that extended beyond the effect solely attributable to sex and age. Measures of memory performance derived from the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), including immediate recall, delayed recall, and list learning, occupied the uppermost section of the lineage-specific spectrum. In contrast, indicators of visuospatial processing (e.g., picture naming on the RBANS) and numerical abilities (e.g., forward digit span score on the RBANS and backward digit span score on the Montreal Cognitive Assessment (MoCA)) were situated at the lowest end of the lineage-specific spectrum. A prospective neuropsychological evaluation of asymptomatic, middle-aged offspring of AD patients (N=60) has suggested that paternal vs. maternal AD liability differentially impacts memory and verbal-numerical abilities [34]. Subjects with maternal AD risk examined for that study showed deficits in delayed and immediate recall in the Loewenstein-Acevedo Scale for Semantic Interference and Learning, as well as on the vocabulary section of the Wechsler Intelligence Scale-III test, as compared to healthy controls. In contrast, a parallel evaluation of

subjects with paternal AD risk enrolled in that study revealed worse performance than healthy controls on the Mini-Mental State Exam, a clinical screening tool for cognitive impairment. Nevertheless, no substantial differences in cognitive performance were observed when directly comparing subjects with maternal versus paternal AD lineage—possibly hindered by the small number of subjects with paternal AD risk (N=13). In our study, involving one of the most comprehensive and homogenous cohorts of AD at-risk subjects, a considerable lineage effect surfaced in differentiating numerical and visuospatial abilities from delayed and immediate recall. This distinction emerged in relation to the *APOE* subgroup that best represented the generational population—specifically, ε3 homozygotes. This could imply that the influence of maternal vs. paternal AD risk on age-related cognitive decline diverges independently from the risk and protective effects attributed to ε4 and ε2, respectively. These generational risk effects appear to particularly impact cognitive functions, as no lineage effects with comparable strength were observed within the neurosensory and cardiovascular domains. Instead, these domains predominantly display sex biases previously linked to normal aging processes.

Complementing existing findings, our third intermediate phenotype of AD susceptibility (Fig. 3) highlighted the simultaneous influence of AD lineage and sex on cardiovascular and cognitive risk indicators. Our analysis placed cardiovascular and cognitive risk indicators on opposite ends of the AD-liability space. A first risk group of effects (Fig. 3; bottom left corner) highlighted indicators of cardiovascular health characteristics of *APOE* ε2 carriers (e.g., hyperlipidemia, statins intake, diastolic and systolic blood pressure, and weight) as tied to both sex and AD lineage. Despite the assumed protective effect of *APOE* ε2 against dementia risk, carrying an ε2 allele has been associated with elevated risks for cardio- and neuro-vascular disorders by previous research groups [35-39]. More recent work by Savignac and colleagues highlighted sex-specificity on the association between ε2 carriership and engaging in strenuous sports [40]. We now found that *APOE* ε3/ε2 was driving most of the differential effects on cognitive and cardiovascular risk indicators as compared to ε3/ε3 and ε3/ε4 genotypes (Fig. 3). Our findings are consistent with an interplay between the ε2 allele and cardiovascular fitness that depends on sex. At the opposite end of the AD-liability spectrum, we found a second group of driving effect differences (Fig. 3; upper right corner) that regrouped various dimensions of cognitive performance (e.g., list recognition and recall, delayed and immediate memory, and visuospatial

attention) affected in AD. Previous work has revealed that the ε2 allele is typically protective against global cognitive decline and dementia [41-43]. A possibility is that underlying differences in cardiovascular fitness between men and women could enhance ε2-related sex bias in cognitive performance. Indeed, the higher incidence of fatal vascular complications in midlife in men compared to women can potentially result in a selective advantage against late-life disorders [44]. Men who survive to an older age will likely have a healthier vascular system than women [45]. The onset of dementia symptoms appears earlier in women than in men [1]. As such, AD symptomatology, including cognitive deficits, is often more severe in women than men at the time of diagnosis [46]. The different aging trajectories of cognitive and cardiovascular health indicators in men and women appear to specifically hinge on the sex-spectrum of AD variation (Fig. 3). As such, the third intermediate phenotype of AD susceptibility emphasized the antagonistic effects of *APOE* ε2 on cardiovascular health and cognitive abilities. In particular, we located ε2-associated interlocking effects of both the sex of the at-risk offspring and of the AD-affected parent.

Lineage-specific trajectories of the intermediate phenotypes over a 4-year follow-up

We capitalized on the serial assessments of the PREVENT-AD cohort to track progression in AD risk indicators over a 4-year follow-up period. We integrated changes in susceptibility phenotypes across the rich collection of PREVENT-AD risk indicators over the years. We aimed to identify which of the seven broad categories of AD risk factors track most sex- and lineage-specific fluctuation in AD risk captured by a given intermediate phenotype over time.

For each subject's first and last follow-up visit, we computed the individual expressions of the three intermediate phenotypes. We then examined how the distinct combinations of PREVENT-AD risk indicators, captured by the three dominant intermediate phenotypes, were differentially expressed with regard to sex and AD lineage — analogous to the procedure conducted on the whole set of subject visits. For each of the seven categories of risk indicators, we averaged over the sex-specific and lineage-specific differences discerned by a given intermediate phenotype separately at the first and last visits. In doing so, we quantified the changes over the years in the sex- and lineage-specific AD progression trajectories, as captured by the three intermediate AD phenotypes.

The sex-specific trajectory captured by the first intermediate phenotype of AD susceptibility increased with regards to the cardiovascular domain (Fig. 1). By construction, the first intermediate phenotype captures most of the variance in the PREVENT-AD phenome that can be attributable to *APOE* genotypes. We found that polymorphic variants on the *BDNF* and *HMGCR* genes primarily accounted for this first relationship. As discussed above, HMGCR and APOE are both involved in maintaining cholesterol homeostasis within the central nervous system [25, 26]. It is known from stroke epidemiology that cardiovascular risk factors, such as brain infarction and intracerebral hemorrhage, are more commonly observed in males [33]. Yet, their effects appear to be more severe in females [33]. The heightened vulnerability of females to severe cardiovascular events could potentially be attributed to the complex interplay between cholesterol homeostasis and hormonal changes in the aging brain. HMGCR is a precursor of steroid hormones, including glucocorticoids and estrogen [25]. Estrogen interacts with HMCGR functions through various mechanisms, including changes in gene expression and signal transduction pathways [47]. Brain regions showing a pronounced reduction in cholesterol synthesis with aging, such as the HC, also display a significant concentration of steroid receptors, highlighting a potential overlap between cholesterol regulation and steroid signaling [48, 49]. Hence, the amplified sex biases in the trajectories of cardiovascular indicators might originate from age-related shifts in estrogen availability in the aging brain, particularly in post-menopausal females. Hormonal fluctuations potentially influence downstream cholesterol synthesis, heightening the risk of cholesterol-related pathologies in sex-specific ways.

The second intermediate phenotype of AD susceptibility tracked variation in cognitive risk indicators that depended on AD lineage. The persistent variability of lineage-specific differences in cognitive risk indicators over time could be related to epigenetic inheritance and its effect on DNA methylation in older age [50]. Chromatin remodelling via histone modifications has been identified as a plausible epigenetic mechanism involved in age-related cognitive decline and neurodegenerative diseases [51]. While the specific genes that undergo epigenetic changes in the adult human brain remain elusive, murine models have singled out several markers of memory and cognition as potential targets [51]. Contextual fear learning was shown to induce *bdnf* DNA methylation in the adult rat hippocampus [52]. Epigenetic regulation was a necessary component

of learning as NMDA receptor blockade, which prevented *bdnf* DNA methylation, resulted in a deficit in memory consolidation [52]. Experiencing life stressors can influence offspring's vulnerability to many diseases by provoking a cascade of epigenetic alterations that alter gene expressions over generations [53]. This process, referred to as epigenetic inheritance [50], represents a plausible avenue by which maternal vs. paternal AD lineage could influence cognitive decline over time. The time trajectory in cognitive risk associated with AD lineage could perhaps reflect such underlying effects of epigenetic inheritance on cognition and memory.

The third intermediate phenotype of AD susceptibility exhibited the most pronounced sex bias in the trajectory of neurosensory processing over time (Fig. 3). Indicators of hearing impairment were most different on the sex-spectrum of disease variation (Fig. 3; right radar plots). Several studies have suggested that hearing loss, especially untreated or severe cases, may contribute to cognitive decline and the onset of dementia [54]. Social isolation and communication difficulties due to hearing loss could potentially increase the risk of developing dementia [55]. The quantity and depth of social interactions are thought to largely differ between males and females [56]. In many cases, as a heterosexual couple ages, it is often observed that the female partner takes on a significant role in maintaining and nurturing the social circle [57]. This phenomenon can be attributed to various factors, including societal norms, caregiving responsibilities, and personal preferences. Hence, it is conceivable that age-related hearing loss might have more severe consequences in females, owing to certain societal and biological factors that render them more prone to seek social interaction. In fact, the relationship between hearing loss and social isolation amongst older adults was found to be stronger for females than males [58, 59]. The escalating sex disparities observed in neurosensory processing trajectories over time may be attributed to varying patterns of social engagement between males and females, which are distinctly influenced by the aging process.

Our longitudinal assessment of the three intermediate phenotypes of AD susceptibility has thus singled out cardiovascular and neurocognitive abilities as preferentially affected throughout AD progression as compared to other risk modalities. The first intermediate phenotype underlined cardiovascular health differences between males and females, possibly modulated by fluctuation in estrogen. The second intermediate phenotype of AD susceptibility tracked lineage-specific

variation in cognitive abilities, which could be related to epigenetic inheritance. Finally, the third intermediate phenotype underscored sex disparities in the trajectory of neurosensory processing, with a particular emphasis on hearing impairment. We have thus shown three distinct longitudinal trajectories associated with AD lineage and sex that had separable ties to cardiovascular health, cognitive performance, and neurosensory processing over the years.

Matri- and patrilinear Alzheimer's disease risk is reflected in anatomical subregions in hippocampus and default network

We next assessed whether matri- vs. patrilinear AD risk differentially targeted brain structures known to be broadly vulnerable in AD patients. In a previous UK Biobank study, we identified AD population patterns of structural covariation in HC and DN subregions as a function of AD family risk in ~40,000 participants [40]. Here we aim to identify which of the same 38 HC and 91 DN target subregions show statistically relevant structural deviation regarding matri- vs. patrilinear AD risk in PREVENT-AD participants. As a first cursory analysis, we estimated classification models separately in males and females to dissociate the type of AD lineage (maternal=1, paternal=0) as a function of grey matter volumes in either the 38 HC or 91 DN subregions. We aimed to quantify how much a given micro-anatomically defined subregion reflects maternal vs. paternal AD liability in the context of the whole set of 38 HC subregions or 91 DN subregions.

In the brain analyses on HC subregions, we identified both cis- (father-to-son, mother-to-daughter) and trans- (father-to-daughter, mother-to-son) generational effects of maternal and paternal AD risk (Fig. 4, top plots). We found an apparent lateralization effect of AD lineage in males for the hippocampus. All patrilinear effects were located in the left hemisphere, whereas all matrilinear effects were located in the right hemisphere. Maternal and paternal AD risks were not as clearly distinguished in HC subregions of the female brain. Indeed, all statistically relevant effects were observed on the left hemisphere regardless of the classification outcome. Only one patrilinear effect was found in females in the left presubiculum body. We have thus uncovered male-specific lateralization effects in HC subregions such that patri- vs. matrilinear effects were preferentially found in the left vs. right hemispheres, respectively.

In the analyses on DN subregions, the matri- vs. patrilinear classification model showed considerable sex-biased effects (Fig. 4, bottom plots). Notably, patrilinear effects were strictly identified in the DN in male subjects. Reminiscent of what we found in HC subregions, 4 of the 5 patrilinear effects robust at 80% confidence were found in the left hemisphere. This pattern of cis-generational variation provides a biological ground for father-to-son transmission effects reported by epidemiological reports [60] as possibly rooted in specific architectural brain features. Our data-driven brain-lineage association test supports a male-specific structural association with AD liability most strongly reflected in phylogenetically recent (i.e., allocortical) rather than older (i.e., neocortical) layers of the cortex.

The susceptibility phenotypes interact with HC-DN signatures in estimating matri- vs. patrilinear AD risk

We previously leveraged the wealth of 40,000 UKB MRI visits to derive population signatures of structural co-variation in DN subregions that showed inter-individual variation with microanatomical HC subregions [40]. We are now interested in carrying over our UKB-derived population signatures of HC-DN co-variation in an AD-at-risk cohort. We aimed to assess whether our signatures of HC-DN co-variation, derived from a representative UK population, can successfully track targeted AD markers closely monitored in PREVENT-AD subjects. As a first step, we carried over this knowledge by computing the subject-specific presence of the 25 modes of HC-DN co-variation in the PREVENT-AD cohort (cf. above). Each participant visit was thus supplemented by the expression levels of the 25 HC and 25 DN patterns of structural co-variation corresponding to the 25 pairs of HC-DN co-variation signatures.

We next tested whether we could further characterize the PREVENT-AD-derived intermediate phenotypes by leveraging the UKB-derived population signatures of HC-DN co-variation. Separately for males and females, we built logistic regression models to classify PREVENT-AD at-risk participants with maternal vs. paternal AD lineage (maternal=1, paternal=0). For each sex, we built separate classification models for each of the 25 HC and 25 DN canonical variates, yielding 50 estimated models per sex. The explanatory variables of each

classification model consisted of the three intermediate phenotypes (computed on the whole PREVENT-AD cohort), a given HC or DN co-variation pattern, and the interaction of the given HC or DN co-variation pattern with each of the three intermediate phenotypes. 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 report strong coefficients statistically different from their respective null distribution at 95% confidence.

We found that HC-DN co-variation patterns explained brain variation related to AD family lineage in sex-specific ways. HC-DN covariation patterns contributed to the classification of matri- vs. patrilinear AD risk in males more than in females, reminiscent of the diagnostic test. In contrast, the intermediate phenotypes were of relatively greater importance in females than in males to differentiate between maternal vs. paternal AD risk. The first intermediate phenotype of AD susceptibility showed a robust main effect in females associated with maternal lineage in most models. In most models, the third intermediate phenotype showed the opposite effect in females and was associated with paternal lineage. Most robust interaction effects between the first intermediate phenotype and HC and DN co-variation expressions were found in females and were specific to maternal risk. In contrast, we found robust interaction effects in males for the third intermediate phenotype associated with paternal risk. The second intermediate phenotype showed a unique pattern of sex-specific differentiation in HC and DN subregions. Brain-phenotype interaction effects were more robust in DN subregions for males and in HC subregions for females. Overall, our analysis of maternal vs. paternal AD risk revealed that the intermediate phenotypes derived from the PREVENT-AD cohort had more robust direct effects in explaining variance in the type of AD lineage in females than in males. In contrast, the HC-DN co-variation regimes were initially derived from the UK Biobank imaging cohort and differentiated between maternal and paternal AD lineage in males more than in females. Bridging across these two independently collected large cohorts allowed us to partition the variance in AD familial risk linked to the brain and the phenome. In doing so, we have provided unprecedented evidence of distinct, non-overlapping patterns of structural variation in neocortical subregions of the DN as being jointly tied to male sex and paternal AD liability.

Cis- and trans-generational AD risk is manifested in default network and hippocampus subregions

We next sought to quantify how matri- vs. patrilinear AD risk captured by the PREVENT-AD-derived intermediate phenotypes were expressed in AD-vulnerable brain structures. Our goal was to determine which of the original 38 HC and 91 DN subregions were driving most of the associations between the UKB-derived HC-DN co-variation signatures and the PREVENT-AD-derived intermediate phenotypes of AD susceptibility. We built on the classification models (cf. above) to closely dissect the relationship between the three intermediate phenotypes and HC-DN co-variation signatures encapsulated by the interaction terms. In males and females separately, we multiplied the interaction terms between a given intermediate phenotype and a given HC or DN pattern by the subject-specific expressions of that same pattern. For each subject visit, we obtained 50 brain-phenotype association terms corresponding to the original pairs of 25 HC-DN brain signatures pooled across the variance in AD lineage explained by a giving intermediate phenotype. We projected back the 50 brain-phenotype association terms onto brain space by multiplying them with the respective 38 HC and 91 DN loadings of the original UKB-derived CCA model (c.f. methods). In doing so, we were able to assess the individual contribution of each of the 38 HC and 91 DN subregions of the original allocortical and neocortical atlas to brain-phenotype associations between the three intermediate phenotypes of AD susceptibility and the 25 brain signatures of HC-DN co-variation.

The combined interaction effects of the 25 brain signatures of HC-DN co-variation with the first intermediate phenotype of AD susceptibility revealed sex bias in neocortical and allocortical subregion volumes (Fig. 6, left panels). In females, maternal AD risk highlighted variation in the DN's lateral structures, most pronounced in the left superior temporal gyrus (STG) and right orbitofrontal cortex (OFC). The opposite spatial distribution stood out in males; maternal AD risk was associated with structural variation in the DN's medial structures, notably the bilateral dorsomedial prefrontal cortex (dmPFC) and left precuneus (PCu). In contrast to the spatial separation of DN effects between males and females, a considerable overlap appeared in HC subregions. Notably, the left CA3 head was associated with paternal lineage in both sexes, whereas the left CA4 head, right CA1 body, and right subiculum body and head were associated with

maternal lineage in both sexes. The association strength of these non-linear effects was noticeably smaller in males, which echoes the permutation results from the sex-specific classification analyses. All but one of the 11 brain-phenotype associations that survived the permutation test for the first intermediate phenotype of AD susceptibility were found in female subjects (Fig. 5). The spatial distribution of matri- and patrilinear effects was consistent with the existence of distinct, non-overlapping, anatomical connections between the HC and DN. Indeed, maternal AD risk was associated with variation along HC and DN subregions structurally connected via the fornix white matter tracts. The fornix, which carries fibre bundle axons from the CA1 and subiculum subregions, propagates the only hippocampal output signals that directly go to the orbitofrontal cortex of the DN [61, 62]. This pathway is mainly involved in spatial memory and navigation [63, 64]. The cingulum bundle could represent an alternative route by which HC signals are conveyed to posterior midline structures of the DN to support pattern separation [64]. Pattern completion and separation could heavily rely on a dynamical system sustained by the DG and CA3, coupled with midline structures of the DN [64, 65]. The spatial distribution of patrilinear effects was reminiscent of anatomical connections between CA3 and posterior midline structures of the DN and contrasted with maternal effects found along the fornix tracts. Thus, we have singled out two functionally distinct cortical fibre bundles, the fornix and cingulum, as potential sources of the lineage differentiation captured by the first intermediate phenotype.

The second intermediate phenotype of AD susceptibility was characterized by trans-generational sex effects in neocortical subregions of the DN (Fig. 6, central panels). In females, we found that structural variation in all but two of the 91 DN subregions were related to paternal AD risk. Patrilinear effects were powerful in lateral structures of the DN, notably in the left superior parietal lobule (SPL), left medial temporal gyrus (MTG), right supramarginal gyrus (SMG), and right ventrolateral prefrontal cortex (vlPFC). We found the opposite generational effects in males; structural variation in all but five of the 91 DN subregions was related to maternal lineage. In contrast to what we found in females, matrilinear effects in males were found in both lateral (e.g., left STG) and medial structures (e.g., right PCu, right dmPFC) of the DN. Considerable lateralization emerged in HC subregions for both males and females. The molecular layer and entire DN highlighted patrilinear effects in females and matrilinear effects in males. We found the opposite effect for the left CA4 head, in which we located matrilinear effects in females

and patrilinear effects in males. The ML comprises axons branching from the pyramidal cells of the CA subfields and subiculum [65]. On the posterior part of the HC, the ML merged into the fornix's fimbria, which supplies only direct hippocampal output signals to the DN [61, 62]. The congruence of the matri- and patrilinear effects, found respectively in the ML and entire DN of male and female subjects, is consistent with the fornix's prominent role in sustaining HC-DN communication. Previous research has singled out the fornix fibres among 48 anatomical tracts as most strongly related to DN gray matter patterns [66]. CA4, also known as the hilar region of the dentate gyrus, has the lowest density of pyramidal cells of the cornu Ammonis subfields [67]. In addition, CA4 does not subserve temporal lobe cortical projections, as is the case for the CA1 and presubiculum [68]. CA4's contribution to HC efferent signals is thus of limited reach and presumably bounded to local modulatory control. The sparse communication between the CA4 and neocortex perhaps explains the incongruence of our reported matri- and patrilinear effects. The combined interaction effects of the 25 brain signatures of HC-DN co-variation with the second intermediate phenotype of AD susceptibility thus highlight the ML and, by extension, the fornix in sustaining HC-DN coupling.

The third intermediate phenotype of AD susceptibility revealed cis-generational sex effects in neocortical subregions of the DN (Fig. 6, right panels). While the individual contribution of the HC-DN signatures to brain-phenotype associations was only robust in males (Fig. 5), their combined influence revealed matrilinear effects in females and patrilinear effects in males. Structural variation was most strongly associated with maternal AD risk in females in the left retrosplenial cortex (RSC), left vmPFC, and right angular gyrus (AG). In contrast, patrilinear effects were most prominent in males in the left dmPFC and right temporo-parietal junction (TPJ). We found a significant degree of overlap in HC subregions. Although the magnitude was more substantial for male subjects, we found matrilinear effects in both sexes in the fimbria and CA3 head. We located corresponding matrilinear effects in the left vmPFC of both sexes—the same hemisphere that showed the most robust positive fimbria weights. The fimbria's projections to the vmPFC via the fornix are considered unilateral [61]. We identified congruent matrilinear effects in the vmPFC and fimbria, consistent with the prominent role of the fornix in HC-DN coupling. In males, these matrilinear effects persisted even in the presence of strong patrilinear weights in dorsal parts of the DN (e.g., dmPFC and TPJ). Our classification models for male subjects

highlighted the contrasting presence of matrilinear HC weights and patrilinear DN weights (Fig. 5). Indeed, brain-phenotype associations effects that survived the 95% permutation test in males were of positive signs on the HC models (i.e., associated with maternal lineage), and negative signs (i.e., associated with paternal lineage) on the DN models. Our structural dissection of brain-phenotypes associations on the third intermediate phenotype of AD susceptibility has thus provided an anatomical basis to the contrasting matrilinear HC weights and patrilinear DN weights found in male-specific models. We established that patrilinear weights were mostly located on superficial parts of the DN. In contrast, we located some matrilinear weights in the ventral parts of the DN with known anatomical connections to the HC. Thus, we have found biologically grounded differentiation of AD lineage in HC and DN subregions reminiscent of the unilateral projections from the HC to the DN.

Discussion

We have isolated and characterized a rich collection of matri- and patrilinear effects in *APOE*-related AD risk transmission from one generation to the next. This analysis was critically enabled by the, to our knowledge, richest homogenously acquired prospective cohort of first-degree relatives of AD patients, with >200 carefully curated phenome markers – the PREVENT-AD initiative. In doing so, we uncovered intermediate phenotypes of AD susceptibility that surfaced differently in male vs. female offspring. Concomitantly, the derived intermediate phenotypes were strongly dependent on the sex of the diagnosed parent. Lineage-specific differentiation in the phenome and brain structure became apparent in asymptomatic children of AD patients. Cognitive and cardiovascular risk indicators were most divergent on the lineage spectrum of disease variation compared to other consequences of AD family burden.

These phenome traits were most evident in the associations with the putatively protective *APOE* ε2 carriership and placed global cognitive performance in direct opposition to cardiovascular risk indicators. Paternal and maternal inheritance routes showed distinct biological footprints in the HC and DN of at-risk subjects. Our findings pointed to HC and DN subregions, which are known to be anatomically connected via the fornix fibre pathway [66], as showing the most structural variation with regard to AD lineage. Over the past two decades, several hypotheses have emerged as target candidates in explaining the differential impact of maternal vs. paternal AD liability [69]. We will address the plausible primary sources of biases in AD transmission that could explain the cross-generational effects captured by our derived intermediate phenotypes. These biological mechanisms, potentially at play in our findings, can be regrouped into three kinds of candidate explanations: mitochondrial alteration, epigenetic imprinting, and chromosome X-mediated transmission.

Mitochondrial DNA (mtDNA) is categorically inherited from the mother, making it a biological source of sex bias that can be transmitted over generations [70]. Our first intermediate phenotype highlighted lineage-sex biases for two genetic markers that have previously shown to control mitochondrial functioning: *HMGCR* and *BDNF* polymorphisms. The Met allele of the *BDNF* gene has been associated with decreased in-vivo levels of N-acetyl aspartate, a marker of

mitochondrial oxidative stress, in the hippocampus of human subjects [21]. Similarly, a high dose of HMGCR inhibitors (statins) has been linked to mitochondrial dysfunction and intracellular oxidative stress in human cell cultures [71]. With almost 20 years of standing in AD research, the mitochondrial cascade hypothesis has challenged the brain-centric view of AD by shifting the object of focus to more systemic biochemical features, given that mitochondria are pervasive human body cells [72, 73]. In its original form, the mitochondrial cascade hypothesis holds that mitochondrial dysfunction precedes and potentiates deposition of β-amyloid (Aβ) aggregates [73]. In more recent years, evidence that Aβ alters mitochondrial functioning has informed and refined the original hypothesis [74]. In-vitro analyses have suggested that $A\beta_{42}$ peptides could act as a neurotoxin to induce oxidative stress, which impairs mitochondrial functioning [75]. Inversely, cells expressing mitochondrial DNA from AD subjects have shown elevated oxidative stress markers that, in turn, promoted Aβ toxicity and programmed cell death [76]. Missense mitochondrial genome mutations that might lead to oxidative phosphorylation have been found at a higher rate in AD patients and children of affected mothers [77]. In a female mouse AD model, embryonic hippocampal neurons showed decreased mitochondrial respiration sustained through the reproductive period and most apparent during reproductive decline [78]. Mitochondrial amyloid load in hippocampal CA1 neurons also increased in this female mouse model, echoing findings from human subjects [78]. Indeed, subjects with maternal AD risk have shown elevated CSF levels of F2-isoprostane, a marker of oxidative stress, that co-occurred with lower CSF levels of $A\beta_{42/40}$, a marker of amyloid aggregation [13]. Estrogen-signalling pathways could further enhance sex biases in mitochondrial functions, particularly in hippocampal neurons [79]. In fact, the intact hippocampus has a relatively elevated concentration of steroid hormone receptors compared to other brain regions [49]. Estradiol may promote mitochondrial respiration, ATP generation, and antioxidant mechanisms [80]. Maternally transmitted mitochondrial DNA could provide a genetic ground by which protective or deleterious haplotypes are passed on from generation to generation. The influence of estrogens throughout females' reproductive age up to senescence could further exacerbate lineage biases in mitochondrial functioning by precipitating Aβ aggregates accumulation and AD pathogenesis. Our discovered lineage-sex biases in *HMGCR* and *BDNF* polymorphisms could thus reflect maternally transmitted genomic effects on mitochondrial functioning and be further modulated by related sex hormone pathways.

Another plausible avenue by which the genealogy of AD risk may be contribute to our results hinges on epigenetic modulation. Epigenetic inheritance (modifications of the tails of histones that carry DNA) refers to the phenomenon by which experiences and environmental changes in the parent generation are transmitted across several offspring generations [53]. It can be seen as part of an overarching class of imprinting mechanisms by which parental genome features impact the vulnerability of offspring to many pathological conditions. Epigenetic inheritance can occur in a sex-dependant manner, affecting transcriptional patterns in males and females separately and as a function of the parent's sex. Indeed, the mechanisms by which methylation markers are reprogrammed after fertilization differ in time and nature for paternal and maternal DNA [53]. After the lineage-dependent reprogramming, gametes methylation then differs as a function of sex. Oocytes methylation is slower and dynamic, increases progressively until sex maturation, and declines to around 40% [81]. In contrast, gametes methylation is much faster in males and reaches 90% before birth [81]. The compartmentalization of methylation processes into lineage- and sex-dependent mechanisms makes epigenetic transmission a plausible source of maternal vs. paternal biases in AD risk [50]. The idea that maternal and paternal genomes contribute equally to gametes at meiosis but entail different phenotypic effects during development was popularized over 30 years ago, notably with regard to late-onset disorders [82]. Such imprinting mechanisms have been suspected for a maternal [83] and paternal history of AD [84]. More recently, a case-control analysis that looked at 93 genes with age-specific expression in the brain revealed a significantly higher number of maternal imprints in late-onset AD cases compared to controls [85]. A gene that can be paternally imprinted in the placenta and fetal brain could nonetheless be maternally imprinted in fibroblasts and lymphocytes [85]. Recent evidence indeed suggests that partitioned heritability for maternal and paternal AD risk differ in different tissues. Maternal AD was notably enriched in the thyroid, pituitary, and esophagus. In contrast, paternal AD risk was most prominent in the cortex. These recent insights on AD lineage suggest that maternal AD risk is more systemic, whereas paternal risk could be specially manifested in the brain. Transcriptional reprogramming during senescence could account for non-Mendelian lineage-specific effects in late-life neurodegenerative disorders such as AD [50, 51, 86]. These effects include but are not limited to parent-of-origin effects [83], differential age at onset between maternal and paternal cases [87], chromatin remodelling [51], and age-dependent epigenetic drift [86]. Parent-of-origin effects refer to the phenomena by which the degree of expressions of certain

genes differs by the sex of the parent from which they are inherited [88]. The differential ages at dementia onset of maternally vs. paternally inherited AD cases could represent an instance of parent-of-origin effects. Sibling pairs of affected mothers showed almost identical age at AD onset, independent of sex [87]. The similarity in age at onset within sibling pairs, which these authors strictly observed amongst maternal, but not paternal, cases, could represent a robust clinical marker of maternal transmission. Nonetheless, evidence from monozygotic twin pairs has shown that essentially identical epigenetic profiles in early life can substantially differ in older age [86, 89]. The acceleration of epigenetic modulations with aging, referred to as age-specific epigenetic drift, could be characteristic of late-onset AD cases [86]. In particular, chromatin remodelling is prevalent during senescence and is thought to have impacts on longevity and age-related cognitive decline [51]. While epigenetic inheritance represents a promising imprinting mechanism by which maternal and paternal AD risk differently affect offspring's risk, more work is needed to identify the specific gene, tissues, and cell types in which these parent-of-origin effects are most evident. For these reasons, epigenetic chromatin remodelling via histone modification could represent a potential therapeutic target for age-related disorders because of its modifiable and highly dynamic nature [50, 90].

Lastly, chromosome X-mediated transmission could represent a contingent source of lineage biases in AD risk. While females generally inherit a chromosome X from both parents, males typically receive a sole chromosome from their mother. Harbouring a second X chromosome could represent a biological advantage in conferring resilience towards cognitive deficits in females. Indeed, RNA sequencing of the human dorsolateral prefrontal cortex linked several X-chromosomal genes to slower cognitive decline in autopsy samples from older females but not males [91]. In contrast, the expression of X genes involved in protein folding was associated with neuropathological tau burden in males but not in females [91]. Animal models of AD pathology corroborated these patterns of X chromosome-mediated vulnerability that is specific to one sex [92]. In fact, neurons from wild-type mice with the XY genotype were more vulnerable to Aβ-induced toxicity when exposed to recombinant $A\beta_{1-42}$ than those of mice with the XX genotype [92]. The X chromosome likely drove this effect, as similar findings were obtained when comparing neurons from mice with XY and X0 chromosomes to mice with XX and XXY chromosomes [92]. A recent genome-wide association study in ~40,000 UK Biobank participants

from the imaging cohort identified two X chromosome clusters of brain-gene associations that statistically differ in males and females [93]. The first X chromosome cluster was linked to metrics of white matter integrity, with peaked effects identified for neurite density in the left superior longitudinal fasciculus [93]. The second cluster highlighted measures of gray versus white matter intensity contrast in limbic and temporal regions identified as belonging to the DN [93]. It thus seems that harbouring only one X chromosome leads to poorer health outcomes, possibly by making the brain more vulnerable to age-related neurotoxicity that eventually precipitates AD pathogenesis. A recent meta-analysis has indeed shown that the male sex, defined as having the XY genotype, increased the risk for death in AD by 62% as compared to the female sex, defined as having the XX genotype [92]. Candidate X genes, such as *Kdm6a*, have emerged from recent translational work from mice to human as protective against cognitive deficits and Aβ toxicity [92]. Nonetheless, the frequency of SNPs mutations on the X chromosome is reduced by half in male participants. Larger, population-based genome studies of gene-trait associations are needed to achieve generalizability beyond hand-selected AD clinical samples.

Conclusion

Our data-informed framework identified parent-of-origin biases in *APOE*-related AD risk in the currently available richest homogenous prospective cohort of AD-at-risk subjects. We extend the widely adopted brain-centric view of AD and offer a complete overview of lineage-sex effects in the whole set of phenotypes available in PREVENT-AD. In doing so, we disentangled a considerable part of the heterogeneous nature of AD risk by establishing intermediate phenotypes that statistically differed as a function of maternal and paternal AD risk. Our careful filtering of the ever-growing collection of AD risk indicators advocates for therapeutic avenues centred around sex-specific pathways in AD transmission. The different manifestation of maternal and paternal AD risk in male and female offspring is consistent with the co-existence of two distinct disease risk categories rooted in separate biological mechanisms.

Methods

Population data source

The PREVENT-AD cohort [14] 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, standard deviation [SD] 5 years). Participants of the PREVENT-AD initiative have undergone extensive annual health and cognitive assessments for up to five 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: ε3/3 (51.2%), ε3/4 (33.1%), ε3/2 (10.5%), ε2/4 (3.0%), ε4/4 (2.1%). Data used in preparation of this article were obtained from the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) program (https://douglas.research.mcgill.ca/stop-ad-centre), data release 6.0. Access to the open data inventory can be found online (https://prevent-alzheimer.net. A complete listing of PREVENT-AD Research Group can be found in the PREVENT-AD database: https://preventad.loris.ca/acknowledgements/acknowledgements.php?date=2023-04-01.

Brain-imaging and preprocessing procedures

Population-based cohorts are ideally suited to tease apart subregion-level variation in AD risk. Advances in automatic segmentation techniques for the human HC using ex vivo brain imaging have allowed for subject-specific parcellations that respect the diversity of distinct subregions identified post-mortem [65]. Capitalizing on these ultra-high resolution segmentations, we previously assessed microstructural alterations of the human HC in a newly detailed way that scaled to ~40,000 UK Biobank (UKB) participants primarily of European genetic ancestry [40]. We described AD-related patterns of structural co-variation in DN subregions, which varied in lockstep with fine-grained HC subregions [40]. Working at a population scale made it possible for us to investigate the effect of rare genotypes on brain structure. Specifically, we were able to characterize the effects of *APOE* ε2 and ε4 on inter-individual expressions of HC-DN co-variation

[40]— something out of reach in traditional brain-imaging studies involving small to medium sample sizes. We are now interested in carrying over our UKB-derived population definitions of HC-DN co-variation to unseen PREVENT-AD participants. In doing so, we aimed to examine whether our HC-DN co-variation signatures successfully track targeted AD markers in persons with a parental history of AD. While the UKB sample was designed to be representative of the general population, the PREVENT-AD cohort was established to monitor pre-symptomatic changes in first-degree relatives of AD patients. This unique population of AD-vulnerable individuals, with an estimated 2-3 fold relative increase in dementia risk [15, 16], allows us to systematically assess the expression of our HC-DN co-variation signatures in the context of AD progression.

The PREVENT-AD resource provides brain imaging scanning (including T1-weighted images) for up to four years of follow-up from 386 participants. Separately for the brain-imaging scans 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. In the same way as in our previous publication, we next parsed the cortex volumes from 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 by the reference atlas (https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal/Parcellations/project_to_individual). Covariates (age, age^2, sex, sex*age, and sex*age^2) were regressed out from each brain-derived grey matter volume measure as part of the deconfounding procedure. Age was determined at the time of recruitment and sex was self-reported. The final brain-imaging sample consisted of 368 participants (261 women and 107 men) with a total of 910 individual visits.

We extracted the same collection of brain-image-derived phenotypes of grey matter morphology as our previous HC-DN covariation study in the UKB [40]. We were thus able to compute the expression of the 25 UKB-derived modes of HC-DN co-variation based on grey matter measurements for the identical sets of 91 DN and 38 HC target subregions. For each PREVENT-AD visit, we obtained the subject-specific expression levels of the 25 HC and 25 DN

patterns of structural co-variation (i.e., *canonical variates*) that capture the inter-individual variation in the 25 pairs of HC-DN co-variation signatures. These derived brain measures were fed into our downstream analyses.

Non-imaging data preprocessing

Of the 910 visits with brain imaging and *APOE* genotypes, 109 came from participants with both maternal and paternal AD lineage and were excluded. An additional 62 visits that came from participants with only sibling history of AD were also excluded. We removed 57 visits from participants with ε4/4 (1 male/7 females) and ε2/4 (6 males/8 females) genotypes because of their limited sample sizes. We next balanced the percentage of males and females with ε3/2, ε3/3, and ε3/4 genotypes by dropping visits from female participants at random. We used the same procedure to balance the percentage of males and females with maternal and paternal AD lineage. A total of 250 visits from female participants were dropped in this procedure. In the final balanced sample, the distribution of *APOE* genotypes was the same in males and females and consisted of ε3/2 at 15%, ε3/4 at 35%, and ε3/3 at 50%. The distribution of maternal and paternal AD lineage in males and females was 34% paternal and 66% maternal. The final sample consisted of 432 participant visits, 182 of which (42%) came from male participants.

Intermediate phenotypes of Alzheimer's disease susceptibility

As the backbone of our analysis workflow, we sought to derive intermediate phenotypes of AD susceptibility that partitioned the phenotypic expression of familial risk as a function of *APOE* genetic background. We capitalized on the rich PREVENT-AD indicators set to capture *APOE*-phenotypes associations across 256 risk indicators from 7 broad risk categories: cardiovascular health, cognition, clinical co-morbidities, demographics, disease progression, genetics, and neurosensory assessments. We designed PLS-regression (PLS-R) models in which the *APOE* genotypes (e.g., ε3/2, ε3/3, ε3/4) were estimated based on these 256 risk indicators. PLS-R was a natural choice of method as it is especially suited to disentangle the variance of a high-dimensional set as a function of a targeted outcome. The explanatory input variable set X was constructed from the PREVENT-AD risk indicators (number of subject visits $\times$ 256 phenotypes).

A parallel outcome variable set Y was constructed from the one-hot-encoded *APOE* genotypes (number of subject visits $\times$ 3 genotypes (e.g., ε3/2, ε3/3, ε3/4)):

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

where m denotes the number of PREVENT-AD phenotypes, and p the number of *APOE* genotypes. PLS-R finds latent variables that model X and simultaneously predict Y. The two sets X and Y are decomposed as the dot product of two matrices that represent the model scores (T, U), and loadings (P, Q), respectively. The decomposition of the original variable sets is obtained as follows:

$$X = TP^T + E$$
$$Y = UQ^T + F$$
$$T = XW^*$$

where T and U are matrices of size $n \times l$, P is a matrix of size $m \times l$, Q is a matrix of size $p \times l$, and E and F are matrices of normally distributed error terms for X and Y, respectively. The number of loadings is denoted by l and determined by the rank of X. Following the principle of linear regression, Y can be estimated as a function of X through the following equation:

$$Y = TQ^T + G$$

where G is a matrix of normally distributed residuals. This equation can be re-expressed as the multiple regression model:

$$Y = XW^*Q^T + F$$
$$B = W^*Q^T$$

where B is a matrix of regression weights, equivalent to the coefficients of a multiple regression model. PLS-R thus find a series of L orthogonal latent variables, i.e., t_l, that have maximal covariance with Y but are uncorrelated to each other. These latent variables are ordered according

to the amount of variance of Y that they explain. Formally, the optimization can be described as follows:

$$t_l = Xw_l, \qquad \text{such that} \qquad cov(t_l, Y) = max$$

The goal of our PLS-R application was to derive *APOE*-driven intermediate phenotypes of AD susceptibility. We focused on the first three latent variables to highlight phenotypic variation in AD risk associated with *APOE* ε3/2, ε3/3, and ε3/4. That way, we obtained three different constellations of PREVENT-AD phenotypes that were pooled across *APOE* genetic backgrounds. This approach allowed us to extract three intermediate phenotypes of AD susceptibility that encapsulated different relationships between the PREVENT-AD risk indicators and the three most prevalent *APOE* genotypes.

Assessment of the lineage-specific and sex-specific variation in intermediate phenotypes of Alzheimer's disease susceptibility

We next systematically explored sex-specific and lineage-specific variation on the derived intermediate phenotypes of AD susceptibility. Each intermediate phenotype captures a different fraction of the phenome-wide variation in AD risk that can be attributed to the *APOE* gene. Our goal was to interrogate whether sex and maternal vs. paternal AD lineage influence the expression of AD risk markers in the three intermediate phenotypes. To do so, we ran parallel analyses in which the analogous intermediate phenotypes of AD susceptibility were derived from males and females, and maternal vs. paternal AD lineage, separately. That way, we were in a position to assess which aspects of the phenome-wide variation in AD risk can be attributable to sex and family lineage.

We first separated male (N=182) and female (N=250) subject visits and built two sex-specific PLS-R models in which the *APOE* genotypes (e.g., ε3/2, ε3/3, ε3/4) were estimated based on the standardized 256 PREVENT-AD risk indicators (c.f. above). The distributions of *APOE* genotypes and maternal vs. paternal AD lineage was previously balanced between male and females (c.f. above). In parallel, we built analogous PLS-R models for the subject visits coming from subjects with maternal (N=284) and paternal (N=148) AD lineage. We aimed to

quantitatively compare the weights of each of the 256 PREVENT-AD risk indicators on the PLS-R models derived from males vs. females and from subjects with maternal vs. paternal AD lineage. Before subtracting the x-loadings of the first 3 PLS-R components of two given subgroups (e.g., males vs females, maternal vs. paternal) we computed the Pearson's correlation coefficients between the y-loadings (i.e., the weight of the three most prevalent *APOE* genotypes) of the first 5 PLS-R components of the subgroups we were aiming to compare. Our objective was to ensure that the PLS-R components were associated with corresponding *APOE* genotypes, meaning that they were biologically comparable. After computing Pearson's correlation matrices, we reordered the components of one of the two comparison subgroups based on the strength of its Pearson's correlation with the other subgroup. After hierarchically matching the components based on this procedure, we subtracted the x-loadings for males and females as well as for maternal and paternal lineages to derive category-specific estimates of the generational difference in AD risk. In doing so, we were able to identify which of the PREVENT-AD risk indicators showed the most sex-specific and lineage-specific variation on a given intermediate phenotype of AD susceptibility.

Longitudinal analysis of lineage- and sex-specific variation in Alzheimer's disease intermediate phenotypes

We subsequently examined how the sex-specific and lineage-specific variation in AD risk captured by the three intermediate phenotypes of AD susceptibility changes over time. We capitalized on the serial assessments of the PREVENT-AD cohort to track category-wise progression in AD risk over a 4-year follow-up period. That way, we were able to identify which of the 7 broad categories of AD risk factors (c.f. above) track most of the sex-specific and lineage-specific fluctuation in AD risk for a given intermediate phenotype.

The 432 subject visits composing the balanced sample (c.f. above) came from 233 different PREVENT-AD participants. Of those participants, 126 (55% females, 65% maternal AD lineage) had at least one follow-up visit over the 4-year period of assessment. The first and last follow-up visits of each of these participants served as our two grouping time variables. We ran analogous sex-specific and lineage-specific PLS-R models (c.f. above) on both time points. We again hierarchically matched the PLS-R components between males vs. females and maternal vs.

paternal AD lineage to compare the expressions of our three intermediate phenotypes in the 4 comparison subgroups. Subsequently, we subtracted the x-loadings for males and females as well as for maternal and paternal lineages at each of the two time points to track variation in AD risk on the three intermediate phenotypes. We then averaged over the differences in x-loadings with regard to sex and AD lineage for each of the 7 categories of risk indicators at both time points: cardiovascular health, cognition, clinical co-morbidities, demographics, disease progression, genetics, and neurosensory assessments. In doing so, we were able to derive category-wise estimates of the sex- and lineage-specific differences in AD progression for each of the three intermediate phenotypes.

Alzheimer's disease lineage is expressed in hippocampus and default network subregions

As our core goal, we aimed to elucidate whether maternal vs. paternal AD lineage is linked to specific structural variation in HC and DN subregions. More specifically, we wanted to quantify how much a given subregion contributes to the classification of maternal vs. paternal AD risk in the context of the whole set of 38 HC subregions or 91 DN subregions. This diagnostic test enabled us to pin down which (if any) microstructurally defined subregions within two cortical systems most affected in AD were related to either maternal vs. paternal AD lineage.

We built classification models (logistic regression) to estimate the type of AD lineage (maternal=1, paternal=0) as a function of grey matter volumes in the 38 hippocampus subregions of the FreeSurfer subcortical atlas (c.f. above). Two separate lineage-classification models were built, one for males and one for females. Each classifier took 38 input variables corresponding to the grey matter volumes in the HC subregions. We employed a resampling procedure to account for differences in the numbers of males vs. females, and individuals with maternal vs. paternal AD lineage, that could affect classification toward to most prevalent classes. Across 1,000 iterations, we have drawn 100 males and 100 females, half of which had a history of AD on their mother's side, while the other half had an history of AD on their father's side. That way, we obtain 1,000 different subsamples where male vs. female sexes and maternal vs. paternal AD lineage were equally represented. At each iteration, we randomly shuffled the true outcome of the classification model (maternal=1, paternal=0) 1,000 times and recomputed the classification weights for each

resampled subject dataset. In so doing, we could empirically derive null distributions for the 38 coefficients of the 1,000 subsamples on which we performed two-tail tests for statistical relevance. The analogous classification and resampling analyses were conducted on the set of 91 default network subregions.

Regression of Alzheimer's disease lineage on the intermediate phenotypes and hippocampus-default network signatures

We next tested whether our UKB-derived signatures of HC-DN co-variation relate to the PREVENT-AD-derived intermediate phenotypes in classifying the type of AD lineage. Capitalizing on these two independently collected datasets allowed us to identify clinically relevant aspects of AD risk robustly tracked by the HC-DN co-variation signatures. We thus aimed to build upon the characterization of our population-derived limbic-cortical regimes by linking them with widely established indicators of dementia progression in presymptomatic individuals. We examined whether we could differentiate the type of AD lineage based on a set of explanatory input variables including i) the three PREVENT-AD-derived intermediate phenotypes, ii) individual variation in expression of 25 HC-DN co-variation patterns (i.e., canonical variates), and iii) the interaction between the three intermediate phenotypes and the 25 co-variation pattern expression strengths. For each sex, we built separate classification models for each of the 25 HC and 25 DN canonical variates, yielding a total of 50 estimated models per sex. In each model, the type of AD lineage (encoded as 0 for paternal and 1 for maternal) was regressed on one HC or DN canonical variate, the three PREVENT-AD-derived intermediate phenotypes, and three interaction terms capturing possible non-linear association between each of the three intermediate phenotypes and the given HC or DN pattern, for a total of 7 regression parameters. We thus obtained a total of 100 logistic model fits that sought to explain variance in the family history of AD as a function of these 7 parameters.

As a complementary analysis integrating across the obtained classification models, we performed a rigorous permutation analysis to assess the robustness of each of the 7 classification coefficients. In 1,000 iterations, we randomly shuffled the type of AD lineage (maternal=1, paternal=0) across the 910 participant assessments. We recomputed the otherwise identical

classification models based on the data with randomized outcomes. We recorded the classification coefficients from each of the 1,000 iterations and used them to build empirical null distributions which provided the basis to perform two-tail statistical tests at a 95% confidence level.

Back projection of brain-phenotype interaction onto hippocampus and default network subregions

We next sought to quantify how matri- vs. patrilinear AD risk captured by the PREVENT-AD-derived intermediate phenotypes were expressed in AD-vulnerable brain structures. We used the interaction terms from the 100 classification models (c.f. above) to derived estimates of brain-phenotype associations pooled across AD lineage. The interaction terms encapsulate how much a given HC or DN co-variation patterns is linked to a given intermediate phenotype of AD susceptibility in the context of matri- vs. patrilinear AD risk. We could then quantify how much of the brain-phenotypic variation captured by a given intermediate phenotype is reflected in anatomically defined HC and DN subregions.

In sex-specific analyses, we multiplied the interaction terms for a given intermediate phenotype of AD susceptibility with the expression levels of the corresponding HC or DN co-variation patterns. In doing so, we obtain 50 brain-phenotype associations terms pooled across AD lineage, corresponding to the pairs of 25 HC-DN brain signatures. For each of the three intermediate phenotypes of AD susceptibility, we next projected back the 50 brain-phenotype associations terms onto brain space by multiplying them with the respective HC and DN loadings of the original UKB-derived canonical correlation analysis (CCA) model (see [40]). In doing so, we were in a position to measure the individual contribution of the original 38 HC and 91 DN subregions used in the UKB-derived CCA model to the variation on the three intermediate phenotypes of AD susceptibility derived from the PREVENT-AD cohort.

Acknowledgements

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).

JP and SV were supported by Canadian Institute on Health Research (CIHR), the Fonds de la Recherche du Québec en Santé (FRQS) and the J.L. Levesque Foundation. Data used in the preparation of this article were obtained from the Pre-symptomatic Evaluation of Novel or Experimental Treatments for Alzheimer's Disease (PREVENT-AD) program data release 6.0. PREVENT-AD was launched in 2011 as a $13.5 million, 7-year public-private partnership using funds provided by McGill University, FRQS, an unrestricted research grant from Pfizer Canada, the J.L. Levesque Foundation, the Douglas Hospital Research Centre and Foundation, the Government of Canada, and the Canada Fund for Innovation. Private sector contributions are facilitated by the Development Office of the McGill University Faculty of Medicine and by the Douglas Hospital Research Centre Foundation (http://www.douglas.qc.ca/).

Conflict of interest

DB is shareholder and advisory board member of MindState Design Labs, USA

.

Figures

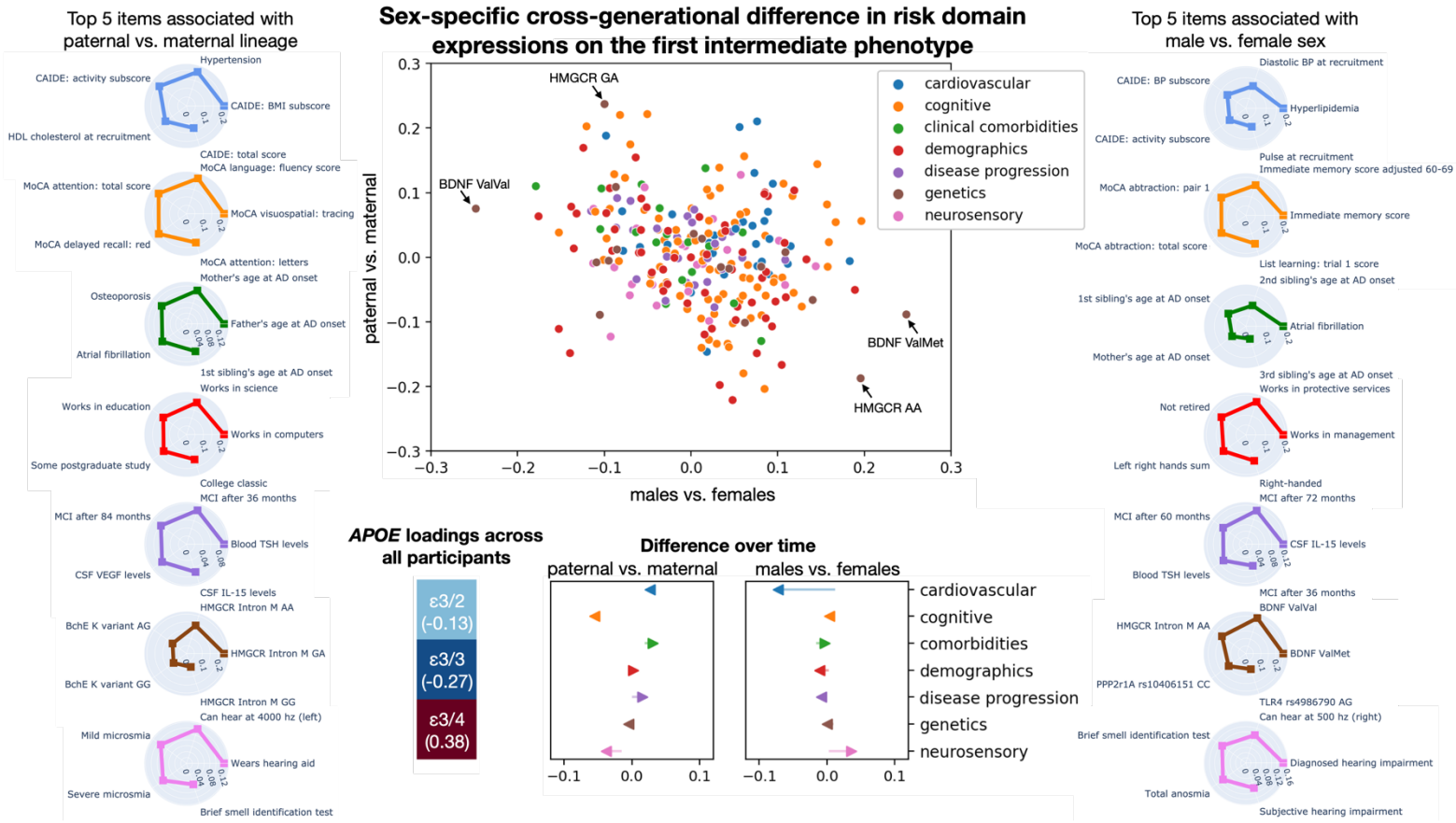


Figure 1. Lineage-specific and sex-biased differences in memory and lipid metabolism tracked by the leading intermediate phenotype of AD susceptibility. We separated male and female at-risk subjects and built two sex-specific PLS regression (PLS-R) models in which the *APOE* genotypes (e.g., ε3/2, ε3/3, ε3/4) were distinguished based on 256 PREVENT-AD phenotypes (dots top middle) spanning 7 broad categories of AD risk factors (colours top middle, cf. legend). Two analogous lineage-specific PLS-R models were built based on at-risk subjects with the conferred maternal and paternal AD liability. We then subtracted the risk indicators model weights for maternal vs. paternal lineage (y-axis) and males vs. females (x-axis). We show the lineage-specific and sex-specific effects against each other for the first PLS-R component (top central panel). The bottom right subplot shows the weights associated with each *APOE* genotype across subjects, while the bottom left subplot shows the mean category-wise difference between maternal vs. paternal lineage and male vs. female sex over time. The accompanying radar plots show the top 5 phenotypes associated with maternal vs. paternal lineage (left) and male vs. female sex (right); one circle for each category of AD risk factors. The first intermediate phenotype of AD susceptibility highlighted cross-generational effects on genetic markers of memory and lipid metabolism. The effect of Val66Met, a missense variant in the gene that codes for BDNF protein, clearly deviated by sex (right genetic radar plot). The effect of SNP rs3846662, which is present in intron M of the HMGCR enzyme, was mostly driven by family lineage (left genetic radar plot). The relative contribution of *APOE* ε4/ε3 to the variation in disease manifestation was stronger than for genotypes ε3/3 and ε3/2. The trajectory of cardiovascular risk indicators showed the most sex bias over time. The first intermediate phenotype of AD susceptibility thus captured cross-generational differences in how proteins are involved in memory and lipid metabolism that persisted through time. BchE: butyrylcholinesterase; BDNF: brain-derived neurotrophic factor, BP: blood pressure; CAIDE: cardiovascular risk factors, aging, and incidence of dementia; CSF: cerebrospinal fluid; HDL: high-density lipoprotein; HMCGR: 3-hydroxy-3-methylglutaryl-CoA reductase; IL-15: interleukin-15; MCI: mild cognitive impairment; MoCA: Montreal cognitive assessment; PPP2r1A: protein phosphatase 2 scaffold subunit Alpha; TLR4: toll-like receptor 4; TSH: thyroid stimulating hormone; VEGF: vascular endothelial growth factor.

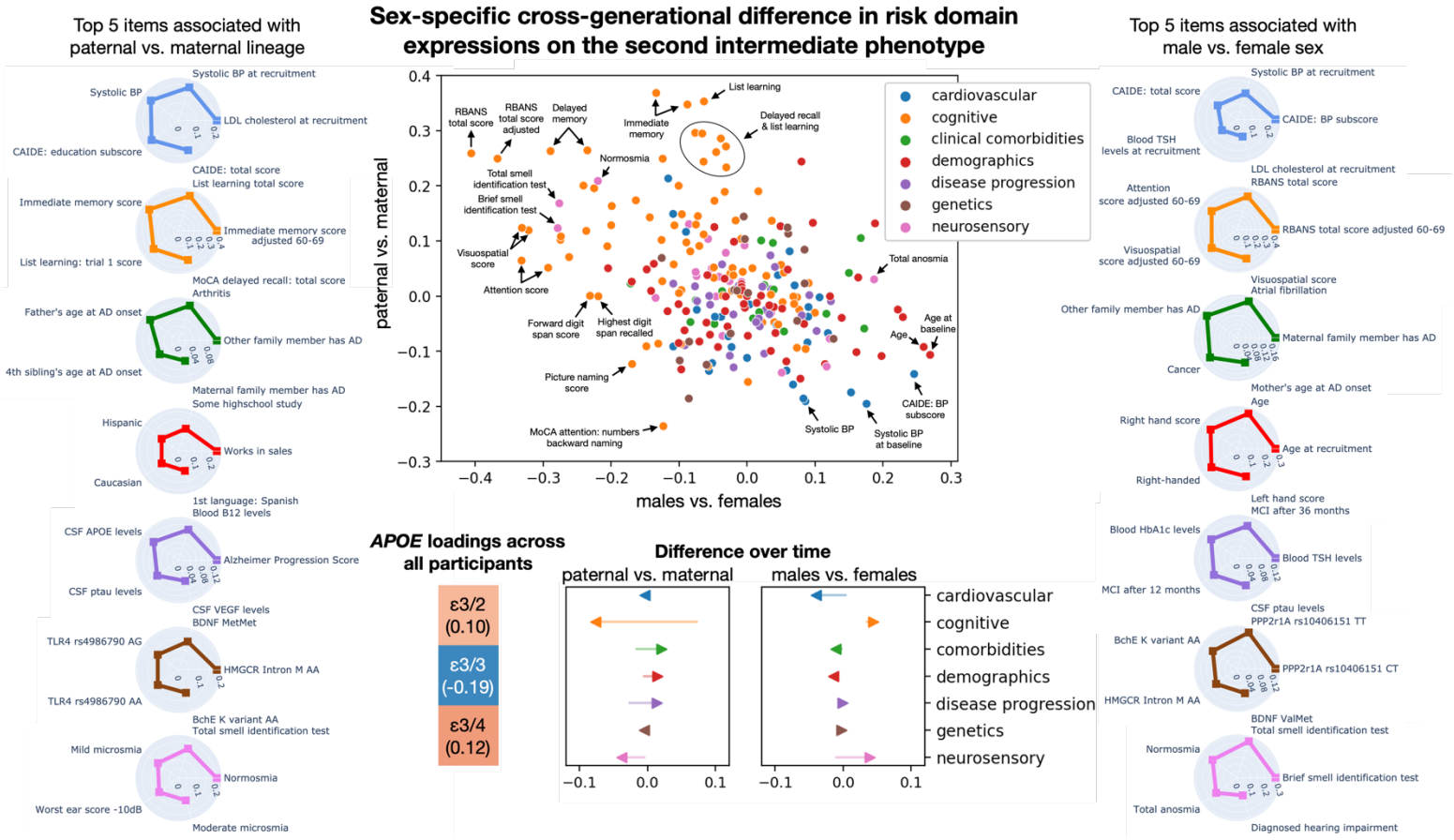


Figure 2. Maternal vs. paternal AD lineage dissociates phonological processes from delayed and immediate recall. We plotted the sex-specific and lineage-specific effects against each other for the second intermediate phenotype of AD susceptibility (top central panel). The bottom right subplot shows the weights associated with each *APOE* genotype across subjects, while the bottom left subplot shows the mean category-wise difference between maternal vs. paternal lineage and male vs. female sex over time. The accompanying radar plots show the top 5 phenotypes associated with maternal vs. paternal lineage (left) and male vs. female sex (right); one circle for each category of AD risk factors. The second intermediate phenotype of AD susceptibility unveiled notable effects of AD lineage on cognitive indicators that extended beyond the effect solely attributable to sex and age. Measures of memory performance and visuospatial attention were situated at the opposite extremes of the lineage-specific spectrum of disease variation. We also revealed a prevailing influence of sex on neurosensory (top right corner) and cardiovascular markers (bottom left corner), accompanied by a comparatively minor effect of AD lineage. The lineage-specific difference in AD risk captured by the cognitive indicators showed the most variation over time. Hence, the second intermediate phenotype of AD susceptibility was thus characterized by a lineage-specific separation of cognitive capacities, distinguishing those related to memory from those linked to numerical and visuospatial abilities. BchE: butyrylcholinesterase; BDNF: brain-derived neurotrophic factor; CAIDE: cardiovascular risk factors, aging, and incidence of dementia; CDK5RAP2: CDK5 regulatory subunit associated protein 2; CSF: cerebrospinal fluid; HbA1c: hemoglobin A1C; HMGCR: 3-hydroxy-3-methylglutaryl-CoA reductase; MCI: mild cognitive impairment; MoCA: Montreal cognitive assessment; PPP2r1A: protein phosphatase 2 scaffold subunit Aalpha; RBANS: repeatable battery for assessment of neuropsychological status; TLR4: toll-like receptor 4; VEGF: vascular endothelial growth factor.

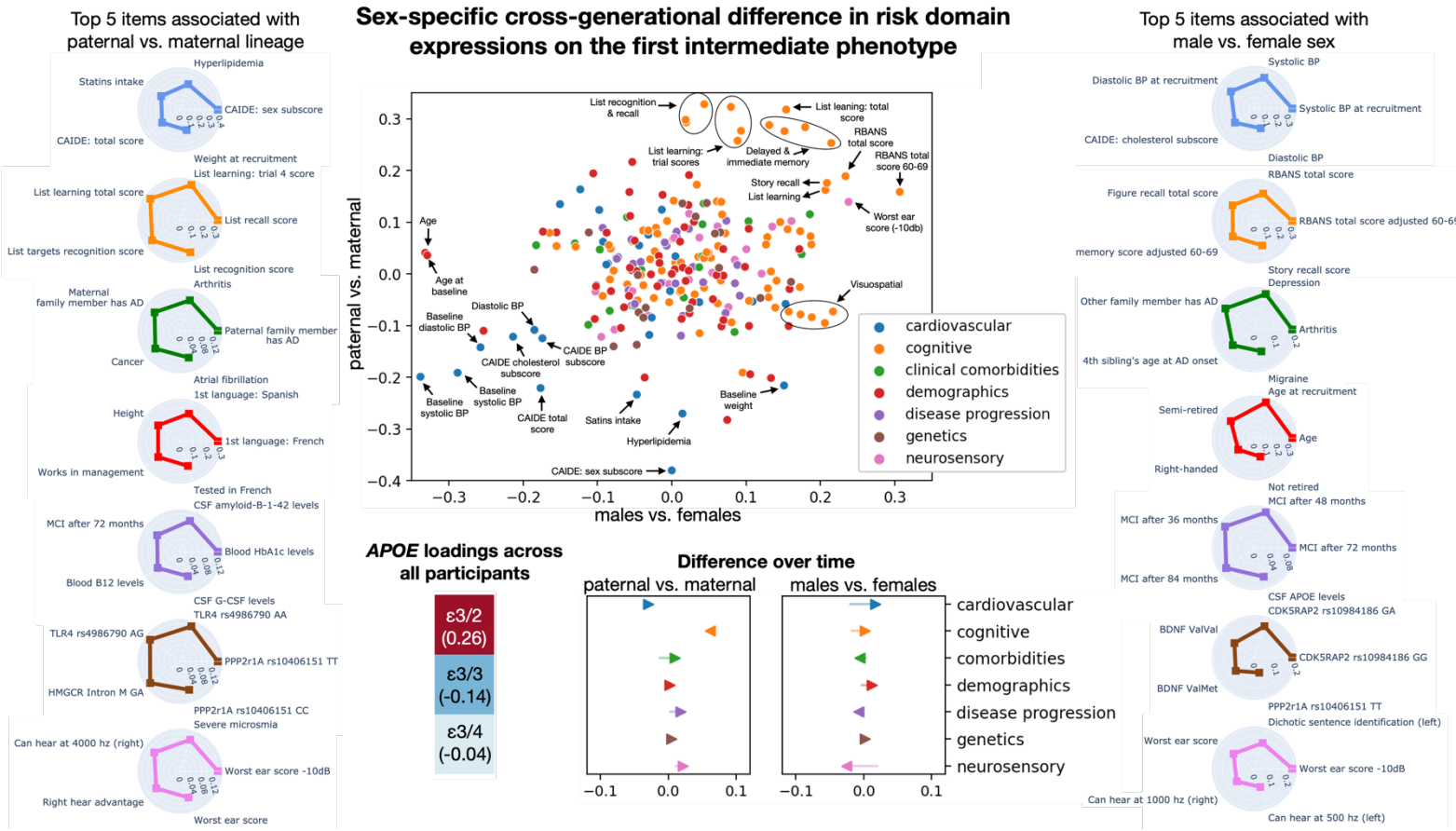


Figure 3. Memory and cardiovascular health outcomes strongly depend on lineage of AD liability transmission. We plotted the lineage-specific and sex-specific effects against each other for the third intermediate phenotype of AD susceptibility (top central panel). The bottom right subplot shows the weights associated with each *APOE* genotype across subjects, while the bottom left subplot shows the mean category-wise difference between maternal vs. paternal lineage and male vs. female sex over time. The accompanying radar plots show the top 5 phenotypes associated with maternal vs. paternal lineage (left) and male vs. female sex (right); one circle for each category of AD risk factors. The third intermediate phenotype of AD susceptibility highlighted the combined influence of AD lineage and sex on cognitive abilities and cardiovascular health. The additive cross-generational effects placed cardiovascular (bottom left corner) and cognitive risk indicators (top right corner) in opposite directions on the AD-liability spectrum, with regard to both sex- and lineage-specific variation. The relative contribution of *APOE* ε3/ε2 to the variation in disease manifestation was stronger than for genotypes ε3/3 and ε3/4. Over time, the sex-specific differences in AD risk captured by the cardiovascular and neurosensory indicators increased in opposite directions. The combined effects of AD lineage and sex therefore placed cognitive and cardiovascular risk indicators at antipode on the AD-liability spectrum derived from the third profile of AD susceptibility. BchE: butyrylcholinesterase; BDNF: brain-derived neurotrophic factor; BP= blood pressure; CAIDE= cardiovascular risk factors, aging, and incidence of dementia; CDK5RAP2: CDK5 regulatory subunit associated protein 2; CSF: cerebrospinal fluid; G-CSF: granulocyte colony-stimulating factor; HMGCR: 3-hydroxy-3-methylglutaryl-CoA reductase; MCI: mild cognitive impairment; MoCA: Montreal cognitive assessment; RBANS: repeatable battery for assessment of neuropsychological status; TLR4: toll-like receptor 4; VEGF: vascular endothelial growth factor.

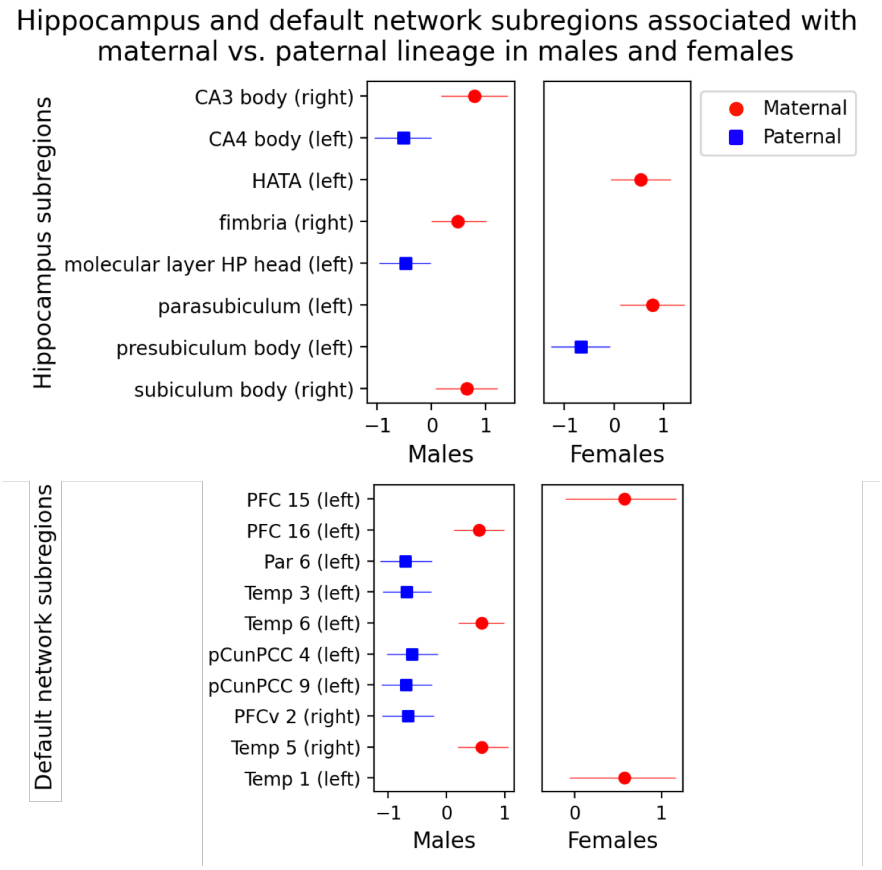


Figure 4. Maternal and paternal Alzheimer's disease liability is tied to specific HC and DN subregions. We built classification models to estimate the type of AD lineage (maternal=1, paternal=0) as a function of grey matter volumes in the set of 38 HC subregions and that of 91 DN subregions. The goal of these analyses was to quantify how much a given subregion contributes to the classification of maternal vs. paternal AD risk in the context of the whole set of either 38 HC subregions or 91 DN subregions. For each atlas definition set, two separate classification models were built, one for males and one for females, that each had 38 or 91 input variables corresponding to the grey matter volumes in HC and DN subregions, respectively. We employed a resampling procedure to account for differences in the derived model effects in males vs. females and individuals with maternal vs. paternal AD lineage. We assessed the robustness of our findings by comparing each of the 38 HC and 91 DN coefficients from the male- and female-specific models to empirically built null distributions obtained through permutation testing. Only the coefficients that were statistically different from their respective null distributions 80% of the time are presented. In the HC models (top panel), we found matrilinear effects located especially to the CA3 body, fimbria, and subiculum body in males, and in the HATA and parasubiculum body in females. We identified patrilinear effects in the CA4 body and molecular layer in males, and presubiculum body in females. All robust patrilinear effects were located to the left hippocampus, regardless of sex. The distribution of matri- and patrilinear effects in DN subregions showed more sex biases than for HC subregions. We identified matrilinear effects in frontal and temporal subregions of the DN that spanned both hemispheres in males and females (bottom panels). In contrast, patrilinear effects were only identified in males and in majority located in the left hemisphere (left bottom panel). No overlaps in matri- and patrilinear effects were observed for any HC or DN subregions. While cis- and trans-generational effects of maternal AD lineage were found in both brain systems, only cis-generational effects of paternal AD risk (i.e., father-to-son) were observed in neocortical subregions of the DN. Our in-depth pattern-learning approach, therefore, detected a male-specific structural association with paternal AD risk reflected in phylogenetically more recent as opposed to older layers of the cortex. HATA: hippocampus–amygdala-transition-area; pCUNPCC: precuneus/posterior cingulate cortex; PFC: prefrontal cortex; PFCv: ventral prefrontal cortex; Par: parietal cortex, Temp: temporal cortex.

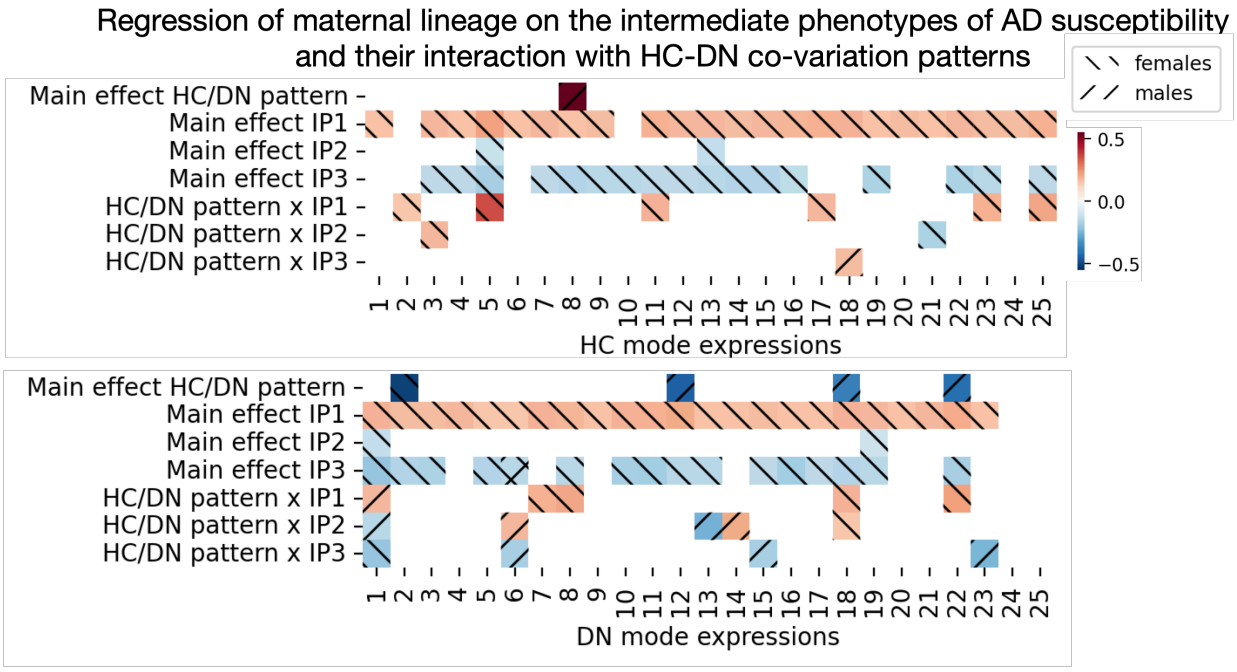


Figure 5. PREVENT-AD phenotypes of AD susceptibility interact with HC-DN population covariation signatures in explaining matri- vs. patrilinear AD risk. We carried over our previously established UKB-derived population brain signatures of HC-DN co-variation (ref. [40]) to the AD-at-risk PREVENT-AD cohort. We assessed whether our HC-DN co-variation signatures, which extracted knowledge from 40,000 UKB MRI visits, can track variation on 256 rich AD markers captured by the three intermediate phenotypes derived from the PREVENT-AD cohort. We computed the expression levels of the 25 UKB-derived modes of HC-DN co-variation in PREVENT-AD subjects based on grey matter measurements in the exact same sets of 91 DN and 38 HC subregions. For each visit, we obtained 25 HC and 25 DN patterns of structural co-variation that correspond to the 25 pairs of HC-DN co-variation signatures. In separated analyses for males and females, we built logistic regression models in which we classified maternal vs. paternal AD lineage (maternal=1, paternal=0) as a function of subject-specific expressions of the 3 intermediate phenotypes (computed on the whole PREVENT-AD cohort), a given HC or DN co-variation pattern, and the interaction of the given HC or DN co-variation pattern with each of the 3 intermediate phenotypes. Each individual classification model is represented as a distinct column on the above heatmaps, on which the HC models (top panel) are separated from the DN models (bottom panel). Effects that were statistically robust in males vs. females at 95% are distinguished with distinct hatching patterns (see legend). In rare cases where an effect was significant in both sexes, hatching patterns were superimposed. The relative contribution of the PREVENT-AD-derived intermediate phenotypes was more important in females than in males. Robust main effects of the first and third intermediate phenotypes were found in females and linked to maternal and paternal AD lineage, respectively. In contrast, the relative contribution of the UKB-derived HC-DN co-variation signatures was more important in males than in females. In males, robust main effects of the HC-DN signatures were found on mode 8 on the HC side, and on modes 12, 18, and 22 on the DN side. A single main effect of mode 2 on the DN side was observed in females. Most interaction effects between the first intermediate phenotype and HC and DN co-variation patterns were found in females and were associated with maternal risk. The opposite was found for intermediate phenotype 3; most brain-phenotype interaction effects were found in males and were associated with paternal risk. We have thus established that the PREVENT-AD-derived intermediate phenotypes are relatively more important in females than in males in driving the classification of maternal vs paternal AD risk. Males rather showed relatively stronger main effects of the UKB-derived HC-DN co-variation signatures, which is consistent with the male-specific constellations of structural associations detected in the diagnostic test. HC: hippocampus; DN: default network; IP: intermediate phenotype of AD susceptibility, UKB: UK Biobank.

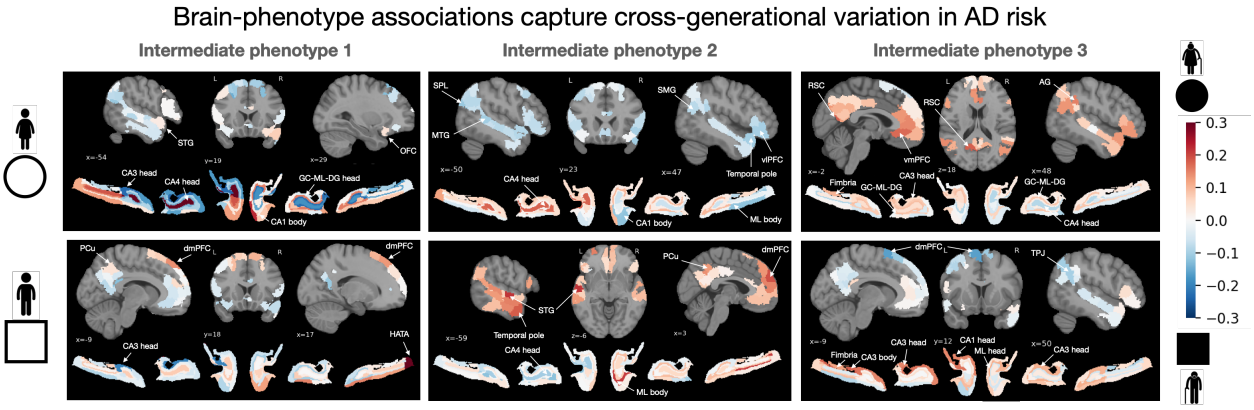


Figure 6. Cis- and trans-generational AD risk manifests in DN and HC subregions. For each of the three intermediate phenotypes of AD susceptibility, we multiplied the interaction terms derived from sex-specific classification models with the subject-specific expressions of the respective HC or DN co-variation pattern. We thus derived 25 new relevance quantities for the HC and DN volume features by participant for each of the three intermediate phenotypes. For each intermediate phenotype, we next mapped back the 25 latent variables for the HC and DN onto the brain by respectively multiplying the HC and DN values with the 38 HC and 91 DN loadings of the canonical correlation analysis (CCA) model trained on the UKB. In doing so, we obtained an expression level of each of the three intermediate phenotypes derived from PREVENT-AD in the original 38 HC and 91 DN subregions used in the CCA model. We plotted the average expression level of the three intermediate phenotypes in each HC and DN subregion in male and female PREVENT-AD subjects. Circles indicate females, squares males. Full shading indicates maternal vs. parental AD cases. The first intermediate phenotype of AD susceptibility revealed sex biases in neocortical and allocortical subregions (left panels). Some degree of spatial heterogeneity in matrilinear effects was found between sexes. Matrilinear weights were located in lateral structures of the DN in females (e.g., left STG, right OFC) as compared to more medial parts of the DN in males (e.g., bilateral dmPFC, left PCu). A considerable degree of spatial overlap was found in allocortical subregions of the HC. The left CA3 head was associated with paternal lineage in both sexes, whereas the left CA4 head, right CA1 body, and right subiculum body and head were associated with maternal lineage in both sexes. The second intermediate phenotype of AD susceptibility was characterized by trans-generational sex effects (i.e., father-to-daughter, mother-to-son) in neocortical subregions of the DN (central panels). Almost the entirety of the DN showed patrilinear weights in females, and matrilinear weights in males. While both the left CA4 and right ML had recurring influence on brain-phenotype associations on both sexes, only the ML and DN weights were of congruent sign. The third intermediate phenotype of AD susceptibility revealed cis-generational sex effects (i.e., mother-to-daughter, father-to-son) in neocortical subregions of the DN (right panels). Matrilinear effects peaked in females in the left RSC, left vmPFC, and right AG. In contrast, patrilinear effects were most prominent in males in the left dmPFC and right TPJ. A significant degree of overlap was observed in HC subregions; matrilinear weights were found in the fimbria and CA3 head in both sexes. Nonetheless, the magnitude of these effects was stronger for male subjects. Our carefully carried structural dissection of brain-phenotype associations has thus grounded cis- and trans-generational effects of AD lineage in HC and DN microstructure. Across intermediate phenotypes, we have highlighted allocortical (e.g., CA1, ML) and neocortical subregions (e.g., vmPFC, OFC) structurally connected via the fornix as being jointly tied to matri- vs. patrilinear AD risk. AG: angular gyrus; dmPFC: dorsomedial prefrontal cortex; GC-ML-DG: granule cell layer and molecular layer of the dentate gyrus; HATA: hippocampus–amygdala-transition-area; ML: molecular layer of the subiculum and CA subfields; MTG: middle temporal gyrus; OFC: orbitofrontal cortex; PCu: precuneus; RSC: retrosplenial cortex; SPL: superior parietal lobule; STG: superior temporal gyrus; TPJ: temporo-parietal junction; vlPFC: ventrolateral prefrontal cortex; vmPFC: ventromedial prefrontal cortex.

References

1. Breitner, J.C., et al., *Familial aggregation in Alzheimer's disease: Comparison of risk among relatives of early-and late-onset cases, and among male and female relatives in successive generations.* Neurology, 1988. **38**(2): p. 207-207.
2. Green, R.C., et al., *Risk of dementia among white and African American relatives of patients with Alzheimer disease.* Jama, 2002. **287**(3): p. 329-336.
3. Association, A.s., *2014 Alzheimer's disease facts and figures.* Alzheimer's & Dementia, 2014. **10**(2): p. e47-e92.
4. Duara, R., et al., *A comparison of familial and sporadic Alzheimer's disease.* Neurology, 1993. **43**(7): p. 1377-1377.
5. Mosconi, L., et al., *Maternal family history of Alzheimer's disease predisposes to reduced brain glucose metabolism.* Proceedings of the National Academy of Sciences, 2007. **104**(48): p. 19067-19072.
6. Honea, R.A., et al., *Reduced gray matter volume in normal adults with a maternal family history of Alzheimer disease.* Neurology, 2010. **74**(2): p. 113.
7. Mosconi, L., et al., *Declining brain glucose metabolism in normal individuals with a maternal history of Alzheimer disease.* Neurology, 2009. **72**(6): p. 513-520.
8. Honea, R.A., et al., *Progressive regional atrophy in normal adults with a maternal history of Alzheimer disease.* Neurology, 2011. **76**(9): p. 822.
9. Cannon-Albright, L.A., et al., *Relative risk for Alzheimer disease based on complete family history.* Neurology, 2019. **92**(15): p. e1745-e1753.
10. Berti, V., et al., *Structural brain changes in normal individuals with a maternal history of Alzheimer's.* Neurobiology of aging, 2011. **32**(12): p. 2325. e17-2325. e26.
11. Mosconi, L., et al., *Reduced mitochondria cytochrome oxidase activity in adult children of mothers with Alzheimer's disease.* Journal of Alzheimer's Disease, 2011. **27**(3): p. 483-490.
12. Okonkwo, O.C., et al., *Cerebral blood flow is diminished in asymptomatic middle-aged adults with maternal history of Alzheimer's disease.* Cerebral cortex, 2014. **24**(4): p. 978-988.
13. Mosconi, L., et al., *Oxidative stress and amyloid-beta pathology in normal individuals with a maternal history of Alzheimer's.* Biological psychiatry, 2010. **68**(10): p. 913-921.
14. Tremblay-Mercier, J., et al., *Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease.* NeuroImage: Clinical, 2021. **31**: p. 102733.
15. Breitner, J.C., et al., *APOE-epsilon4 count predicts age when prevalence of AD increases, then declines: the Cache County Study.* Neurology, 1999. **53**(2): p. 321-31.
16. Huang, W., et al., *APOE genotype, family history of dementia, and Alzheimer disease risk: a 6-year follow-up study.* Arch Neurol, 2004. **61**(12): p. 1930-4.
17. Lu, B., *BDNF and activity-dependent synaptic modulation.* Learning & memory, 2003. **10**(2): p. 86-98.
18. Murer, M.G., Q. Yan, and R. Raisman-Vozari, *Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease.* Progress in neurobiology, 2001. **63**(1): p. 71-124.
19. Phillips, H.S., et al., *BDNF mRNA is decreased in the hippocampus of individuals with Alzheimer's disease.* Neuron, 1991. **7**(5): p. 695-702.

20. Tsai, S.-J., *Critical Issues in BDNF Val66Met Genetic Studies of Neuropsychiatric Disorders.* Frontiers in Molecular Neuroscience, 2018. **11**.
21. Egan, M.F., et al., *The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function.* Cell, 2003. **112**(2): p. 257-269.
22. Fukumoto, N., et al., *Sexually dimorphic effect of the Val66Met polymorphism of BDNF on susceptibility to Alzheimer's disease: New data and meta-analysis.* American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 2010. **153B**(1): p. 235-242.
23. Lin, Y., et al., *Association of rs6265 and rs2030324 polymorphisms in brain-derived neurotrophic factor gene with Alzheimer's disease: a meta-analysis.* PLoS One, 2014. **9**(4): p. e94961.
24. Karlic, H. and F. Varga, *Mevalonate Pathway*, in *Encyclopedia of Cancer (Third Edition)*, P. Boffetta and P. Hainaut, Editors. 2019, Academic Press: Oxford. p. 445-457.
25. Poirier, J., *Apolipoprotein E, cholesterol transport and synthesis in sporadic Alzheimer's disease.* Neurobiology of aging, 2005. **26**(3): p. 355-361.
26. Yu, C., K.L. Youmans, and M.J. LaDu, *Proposed mechanism for lipoprotein remodelling in the brain.* Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 2010. **1801**(8): p. 819-823.
27. Wall, C.T.J., et al., *Mitochondrial respiratory chain dysfunction alters ER sterol sensing and mevalonate pathway activity.* Journal of Biological Chemistry, 2022. **298**(3).
28. Leduc, V., et al., *HMGCR is a genetic modifier for risk, age of onset and MCI conversion to Alzheimer's disease in a three cohorts study.* Molecular Psychiatry, 2015. **20**(7): p. 867-873.
29. Leduc, V., et al., *Role of rs3846662 and HMGCR alternative splicing in statin efficacy and baseline lipid levels in familial hypercholesterolemia.* Pharmacogenet Genomics, 2016. **26**(1): p. 1-11.
30. Leduc, V., et al., *Effects of rs3846662 Variants on HMGCR mRNA and Protein Levels and on Markers of Alzheimer's Disease Pathology.* Journal of Molecular Neuroscience, 2016. **58**(1): p. 109-119.
31. Schiffman, S.S., *Taste and Smell Losses in Normal Aging and Disease.* JAMA, 1997. **278**(16): p. 1357-1362.
32. Strait, J.B. and E.G. Lakatta, *Aging-associated cardiovascular changes and their relationship to heart failure.* Heart failure clinics, 2012. **8**(1): p. 143-164.
33. Appelros, P., B. Stegmayr, and A. Terént, *Sex differences in stroke epidemiology: a systematic review.* Stroke, 2009. **40**(4): p. 1082-1090.
34. Wilson, K.E., et al., *Individual cognitive and depressive traits associated with maternal versus paternal family history of late-onset Alzheimer's disease: proactive semantic interference versus standard neuropsychological assessments.* Personalized Medicine in Psychiatry, 2018. **11**: p. 1-6.
35. Nelson, P.T., et al., *APOE-ε2 and APOE-ε4 Correlate With Increased Amyloid Accumulation in Cerebral Vasculature.* Journal of Neuropathology & Experimental Neurology, 2013. **72**(7): p. 708-715.
36. Yu, L., et al., *APOE and cerebral amyloid angiopathy in community-dwelling older persons.* Neurobiology of Aging, 2015. **36**(11): p. 2946-2953.

37. McKay, G.J., et al., *Evidence of association of APOE with age-related macular degeneration-a pooled analysis of 15 studies.* Human mutation, 2011. **32**(12): p. 1407-1416.
38. Zhao, N., et al., *APOE ε2 is associated with increased tau pathology in primary tauopathy.* Nature communications, 2018. **9**(1): p. 1-11.
39. Ghebremedhin, E., et al., *Argyrophilic grain disease is associated with apolipoprotein E ε2 allele.* Acta neuropathologica, 1998. **96**(3): p. 222-224.
40. Savignac, C., et al., *APOE alleles are associated with sex-specific structural differences in brain regions affected in Alzheimer's disease and related dementia.* PLOS Biology, 2022. **20**(12): p. e3001863.
41. Corder, E., et al., *Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease.* Nature genetics, 1994. **7**(2): p. 180-184.
42. Farrer, L.A., 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): p. 1349-1356.
43. Berlau, D.J., et al., *APOE ε2 is associated with intact cognition but increased Alzheimer pathology in the oldest old.* Neurology, 2009. **72**(9): p. 829-834.
44. Chêne, G., et al., *Gender and incidence of dementia in the Framingham Heart Study from mid-adult life.* Alzheimer's & Dementia, 2015. **11**(3): p. 310-320.
45. Ferretti, M.T., et al., *Sex differences in Alzheimer disease—the gateway to precision medicine.* Nature Reviews Neurology, 2018. **14**(8): p. 457-469.
46. Pradier, C., et al., *The Mini Mental State Examination at the time of Alzheimer's disease and related disorders diagnosis, according to age, education, gender and place of residence: a cross-sectional study among the French National Alzheimer database.* PloS one, 2014. **9**(8): p. e103630.
47. Fester, L., et al., *Cholesterol-promoted synaptogenesis requires the conversion of cholesterol to estradiol in the hippocampus.* Hippocampus, 2009. **19**(8): p. 692-705.
48. Thelen, K.M., et al., *Cholesterol synthesis rate in human hippocampus declines with aging.* Neuroscience Letters, 2006. **403**(1): p. 15-19.
49. Galea, L.A., et al., *Sex, hormones and neurogenesis in the hippocampus: hormonal modulation of neurogenesis and potential functional implications.* Journal of neuroendocrinology, 2013. **25**(11): p. 1039-1061.
50. Bellver-Sanchis, A., M. Pallàs, and C. Griñán-Ferré, *The Contribution of Epigenetic Inheritance Processes on Age-Related Cognitive Decline and Alzheimer's Disease.* Epigenomes, 2021. **5**(2): p. 15.
51. Sen, P., et al., *Epigenetic mechanisms of longevity and aging.* Cell, 2016. **166**(4): p. 822-839.
52. Lubin, F.D., T.L. Roth, and J.D. Sweatt, *Epigenetic regulation of BDNF gene transcription in the consolidation of fear memory.* J Neurosci, 2008. **28**(42): p. 10576-86.
53. Lacal, I. and R. Ventura, *Epigenetic Inheritance: Concepts, Mechanisms and Perspectives.* Frontiers in Molecular Neuroscience, 2018. **11**.
54. Lin, F.R., et al., *Hearing Loss and Incident Dementia.* Archives of Neurology, 2011. **68**(2): p. 214-220.
55. Panza, F., V. Solfrizzi, and G. Logroscino, *Age-related hearing impairment—a risk factor and frailty marker for dementia and AD.* Nature Reviews Neurology, 2015. **11**(3): p. 166-175.

56. David-Barrett, T., et al., *Women Favour Dyadic Relationships, but Men Prefer Clubs: Cross-Cultural Evidence from Social Networking.* PLOS ONE, 2015. **10**(3): p. e0118329.
57. Dykstra, P.A. and T. Fokkema, *Social and Emotional Loneliness Among Divorced and Married Men and Women: Comparing the Deficit and Cognitive Perspectives.* Basic and Applied Social Psychology, 2007. **29**(1): p. 1-12.
58. Shukla, A., et al., *Hearing Loss, Loneliness, and Social Isolation: A Systematic Review.* Otolaryngology–Head and Neck Surgery, 2020. **162**(5): p. 622-633.
59. Mick, P., I. Kawachi, and F.R. Lin, *The Association between Hearing Loss and Social Isolation in Older Adults.* Otolaryngology–Head and Neck Surgery, 2014. **150**(3): p. 378-384.
60. Ehrenkrantz, D., et al., *Genetic epidemiological study of maternal and paternal transmission of Alzheimer's disease.* American journal of medical genetics, 1999. **88**(4): p. 378-382.
61. Aggleton, J.P., et al., *Complementary Patterns of Direct Amygdala and Hippocampal Projections to the Macaque Prefrontal Cortex.* Cerebral Cortex, 2015. **25**(11): p. 4351-4373.
62. Barbas, H. and G.J. Blatt, *Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey.* Hippocampus, 1995. **5**(6): p. 511-533.
63. O'keefe, J. and L. Nadel, *The hippocampus as a cognitive map*. 1978: Oxford university press.
64. Bubb, E.J., C. Metzler-Baddeley, and J.P. Aggleton, *The cingulum bundle: anatomy, function, and dysfunction.* Neuroscience & Biobehavioral Reviews, 2018. **92**: p. 104-127.
65. Iglesias, J.E., 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**: p. 117-137.
66. Kernbach, J.M., et al., *Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants.* Proceedings of the National Academy of Sciences, 2018. **115**(48): p. 12295-12300.
67. Mani, R.B., J.B. Lohr, and D.V. Jeste, *Hippocampal pyramidal cells and aging in the human: A quantitative study of neuronal loss in sectors CA1 to CA4.* Experimental Neurology, 1986. **94**(1): p. 29-40.
68. Blatt, G.J. and D.L. Rosene, *Organization of direct hippocampal efferent projections to the cerebral cortex of the rhesus monkey: projections from CA1, prosubiculum, and subiculum to the temporal lobe.* Journal of Comparative Neurology, 1998. **392**(1): p. 92-114.
69. Mattson, M.P., *Mother's legacy: mitochondrial DNA mutations and Alzheimer's disease.* Trends in neurosciences (Regular edition), 1997. **20**(9): p. 373-375.
70. Giles, R.E., et al., *Maternal inheritance of human mitochondrial DNA.* Proceedings of the National academy of Sciences, 1980. **77**(11): p. 6715-6719.
71. Marcheggiani, F., et al., *Modulation of Coenzyme Q10 content and oxidative status in human dermal fibroblasts using HMG-CoA reductase inhibitor over a broad range of concentrations. From mitohormesis to mitochondrial dysfunction and accelerated aging.* Aging (Albany NY), 2019. **11**(9): p. 2565.
72. Swerdlow, R.H. and S.M. Khan, *A "mitochondrial cascade hypothesis" for sporadic Alzheimer's disease.* Medical Hypotheses, 2004. **63**(1): p. 8-20.

73. Swerdlow, R.H., J.M. Burns, and S.M. Khan, *The Alzheimer's disease mitochondrial cascade hypothesis.* Journal of Alzheimer's Disease, 2010. **20**(s2): p. S265-S279.
74. Swerdlow, R.H., *Mitochondria and mitochondrial cascades in Alzheimer's disease.* Journal of Alzheimer's Disease, 2018. **62**(3): p. 1403-1416.
75. Behl, C., et al., *Hydrogen peroxide mediates amyloid β protein toxicity.* Cell, 1994. **77**(6): p. 817-827.
76. Cardoso, S.M., et al., *Mitochondria dysfunction of Alzheimer's disease cybrids enhances Aβ toxicity.* Journal of Neurochemistry, 2004. **89**(6): p. 1417-1426.
77. Davis, R.E., et al., *Mutations in mitochondrial cytochrome c oxidase genes segregate with late-onset Alzheimer disease.* Proceedings of the National Academy of Sciences, 1997. **94**(9): p. 4526-4531.
78. Yao, J., et al., *Mitochondrial bioenergetic deficit precedes Alzheimer's pathology in female mouse model of Alzheimer's disease.* Proceedings of the National Academy of Sciences, 2009. **106**(34): p. 14670-14675.
79. Brinton, R.D., *The healthy cell bias of estrogen action: mitochondrial bioenergetics and neurological implications.* Trends in Neurosciences, 2008. **31**(10): p. 529-537.
80. Klosinski, L.P., et al., *White matter lipids as a ketogenic fuel supply in aging female brain: implications for Alzheimer's disease.* EBioMedicine, 2015. **2**(12): p. 1888-1904.
81. Kobayashi, H., et al., *Contribution of Intragenic DNA Methylation in Mouse Gametic DNA Methylomes to Establish Oocyte-Specific Heritable Marks.* PLOS Genetics, 2012. **8**(1): p. e1002440.
82. Moore, T. and D. Haig, *Genomic imprinting in mammalian development: a parental tug-of-war.* Trends in genetics, 1991. **7**(2): p. 45-49.
83. Bassett, S.S., D. Avramopoulos, and D. Fallin, *Evidence for parent of origin effect in late-onset Alzheimer disease.* American Journal of Medical Genetics, 2002. **114**(6): p. 679-686.
84. Farrer, L.A., et al., *Association of Decreased Paternal Age and Late-Onset Alzheimer's Disease: An Example of Genetic Imprinting?* Archives of Neurology, 1991. **48**(6): p. 599-604.
85. Chaudhry, M., et al., *Genetic variation in imprinted genes is associated with risk of late-onset Alzheimer's disease.* Journal of Alzheimer's Disease, 2015. **44**(3): p. 989-994.
86. Wang, S.-C., B. Oelze, and A. Schumacher, *Age-Specific Epigenetic Drift in Late-Onset Alzheimer's Disease.* PLOS ONE, 2008. **3**(7): p. e2698.
87. Gómez-Tortosa, E., et al., *Variability of age at onset in siblings with familial Alzheimer disease.* Archives of neurology, 2007. **64**(12): p. 1743-1748.
88. Hall, J.G., *Genomic imprinting: review and relevance to human diseases.* Am J Hum Genet, 1990. **46**(5): p. 857-73.
89. Martin, G.M., *Epigenetic drift in aging identical twins.* Proceedings of the National Academy of Sciences, 2005. **102**(30): p. 10413-10414.
90. Griñán-Ferré, C., et al., *Understanding epigenetics in the neurodegeneration of Alzheimer's disease: SAMP8 mouse model.* Journal of Alzheimer's Disease, 2018. **62**(3): p. 943-963.
91. Davis, E.J., et al., *Sex-Specific Association of the X Chromosome With Cognitive Change and Tau Pathology in Aging and Alzheimer Disease.* JAMA Neurology, 2021. **78**(10): p. 1249-1254.

92. Davis, E.J., et al., *A second X chromosome contributes to resilience in a mouse model of Alzheimer's disease.* Science Translational Medicine, 2020. **12**(558): p. eaaz5677.
93. Smith, S.M., et al., *An expanded set of genome-wide association studies of brain imaging phenotypes in UK Biobank.* Nature neuroscience, 2021. **24**(5): p. 737-745.

Discussion

We developed a mission-tailored analytical framework specially designed for disentangling sex biases in the phenome as a function of Alzheimer's disease (AD) family risk and its associated deviations in micro-anatomically defined subregions within the hippocampus (HC) and default network (DN). The extensive scope of the UK Biobank (UKB) allowed us to examine sex effects on less common *APOE* genotypes, including ε2/2, which brain-imaging studies have rarely incorporated. This unique opportunity enabled a robust comparison of brain-imaging outcomes associated with ε2 and ε4, revealing distinct dosage effects on the brain and phenome. Our population neuroscience approach unveiled sex-based interaction effects between *APOE* ε2 and HC-DN signatures, affecting both fixed (e.g., family history of Alzheimer's disease and related dementias) and modifiable (e.g., social engagement, physical activity, neuroticism) risk factors. Surprisingly, no such interaction effects were observed with the most studied *APOE* ε4. Unlike conventional health policies targeting cognitive decline in vulnerable populations, our findings spotlight modifiable risk factors linked to the protective nature of the ε2 variant.

Our lineage-by-sex examination of the phenotypic variability of AD family history in PREVENT-AD confirmed these considerable sex biases in the protective effects linked to ε2 while adding a concomitant generational influence. When assessing PREVENT-AD phenotypes across different *APOE* genetic backgrounds, we found that the combined influence of AD lineage and sex on cognitive and cardiovascular risk indicators was most prominent in connection with the ε2 variant. Precisely, our data-informed framework positioned global cognitive performance and cardiovascular risk indicators at opposite extremes of the lineage-by-sex spectrum of disease variation in relation to the ε2 genetic risk. Distinct neuroanatomical patterns emerged in HC and DN subregions based on paternal and maternal inheritance routes. Notably, subregions known for their anatomical connection via the fornix fibre pathway [54] exhibited the most prominent structural variation with regard to AD lineage, corroborating our population insights from the UKB. The ε2 allele surfaced as associated with the most prominent sex bias in modifiable risk indicators for AD, underscoring the importance of redirecting efforts towards exploring and optimizing the lesser-studied protective factors mediated by *APOE* ε2.

Our examination of population patterns of grey matter variation linked to familial AD risk has revealed sex biases in microstructural subregions of the HC and DN that favored female ε2 homozygotes. The synergy between isoform-specific effects in APOE signalling and estrogen is poised to provide an advantage for female ε2 carriers by potentiating the immune response to amyloid pathology. This effect may be especially pronounced in cortical areas linked to the HC, such as subregions of the DN, which are among the earliest sites of amyloid deposition during the development of AD [44]. Evidence suggests that ε2 carriers display elevated baseline APOE levels in regions such as the HC and frontal cortex when compared to ε4 carriers and ε3 homozygotes [74-77]. The isoform-specific effects associated with the APOE protein might be amplified by homeostatic responses to the accumulation of amyloid-beta (Aβ) in the brain [78]. Within the central nervous system (CNS), local APOE synthesis is primarily sustained by astrocytes and microglia [25]. Amidst neurodegeneration and Aβ deposition, microglia are thought to upregulate APOE expression [78]. Hormonal factors potentially play a role in the immune response against amyloid pathology. Estrogen is indeed believed to enhance both astrocytic and microglial APOE production [79]. This effect might stem from the presence of an estrogen-dependent enhancer in the promoter region of the *APOE* gene [80]. The influence of estrogens on APOE expression is thought to be especially prominent in regions with a high concentration of steroid hormone receptors such as the hippocampus' CA1 [79]. The pyramidal cells of CA1 are also recognized for showing transient dendritic changes in early AD development [12]. Our leading signature of AD susceptibility, derived from the UKB, showed the most pronounced regional alterations in grey matter volume within the hippocampus' CA1 and molecular layer, which comprises the axons of the pyramidal cells of CA1 (Chapter 1, Fig. 1). Moreover, we found that the protective effect of *APOE* ε2 on AD risk was specific to females and modulated by HC-DN co-variation patterns (Chapter 1, Fig. 6). This finding aligns with the notion that estrogens may play a central role in influencing isoform-specific effects associated with the APOE protein, which are bound to favour female ε2 homozygotes.

The distinct lipid profile associated with the *APOE* ε2 allele may contribute to a more favorable lipid environment for neuronal health, thus supporting cognitive preservation even in the presence of amyloid pathology. Nonetheless, the low affinity of the ε2 variant to low-density lipoprotein (LDL) receptors comes with a significant drawback: a reduced capacity to efficiently

facilitate the vascular clearance of very low-density lipoproteins [81]. This characteristic, in turn, limits the effectiveness of the ε2 isoform in mediating the removal of cholesterol metabolites and triglycerides from blood vessels. As the neuroprotective effect of estrogen weakens with older age, women become more vulnerable to neurovascular disorders that can ultimately lead to dementia [82]. Our analysis of brain-behaviour associations across 40,000 UKB participants has indeed identified proxies of cardiovascular health (e.g., water mass, fat-free mass, and weight) as strongly related to the relation between HC-DN co-variation and *APOE* ε2 vs. ε4 dosage in older females (Chapter 1, Fig. 3). Conducting a more focused exploration of classical AD risk factors, we identified robust interactions between HC-DN pattern expressions and participation in strenuous sports unique to female ε2 homozygotes (Chapter 1, Fig. 5). Engaging in physical activity holds the potential to counteract the combined impact of genetic and age-related predispositions to cardiovascular complications. These converging findings underscore the interplay between estrogens and cardiovascular fitness in shaping AD risk, challenging the prevailing notion of the *APOE* gene's fixed association with dementia in females.

Expanding on the connection between cardiovascular health and dementia risk, our leading intermediate phenotype of AD susceptibility derived from PREVENT-AD highlighted a significant contributor to cholesterol synthesis as preferentially linked to AD lineage: the HMGCR enzyme (Chapter 2, Fig. 1). Also known as 3-hydroxy-3-methylglutaryl-coenzyme A reductase, HMGCR plays a central role in the mevalonate pathway—the primary pathway for cholesterol production [83]. The HMGCR enzyme is crucial for ensuring an ample supply of cholesterol to neuronal membranes in both the brain and periphery, thereby contributing to the maintenance of synapses [25, 84, 85]. When intracellular concentration of cholesterol in glia cells rises, the inhibition of HMGCR repressed cholesterol synthesis. In parallel, the synthesis of the APOE protein is induced to facilitate the transfer of cholesterol to the extracellular environment. Being homozygous for the A allele (AA) has been identified as one of the most important and common protective variants for sporadic AD, second only to *APOE* ε2 [86]. The A allele in intron M of the *HMGCR* gene possibly decreases AD risk by acting as a natural statin and repressing cholesterol synthesis in human brain cells [86]. This effect, in turn, could compensate for lower APOE steady-state levels measured in the CNS of ε4 carriers and ε3 homozygotes compared to ε2. Estrogen is thought to interact with HMCGR functions through various mechanisms, including changes in

gene expression and signal transduction pathways [87]. Indeed, HMGCR is a precursor of steroid hormones, including glucocorticoids and estrogen [25]. The synergistic roles of *HMGCR* and *APOE* in regulating cholesterol homeostasis within the CNS present a compelling avenue for therapeutic interventions aimed at mitigating cholesterol imbalance and its downstream impact on neurodegenerative processes. Our analysis of familial AD risk across generations highlights how *HMGCR* and *APOE* play complementary roles amid age-related metabolic changes. This gene-gene interaction could be further accentuated by the decline in estrogen levels during aging, particularly in post-menopausal females.

APOE genetic risk is widely examined in isolation. This approach could stem from the widespread notion that *APOE* plays a quasi "monogenic" role in late-onset AD [88]. Nevertheless, by employing a totally data-driven approach, we highlighted gene-gene interactions as central to AD familial risk across hundreds of biological markers. In addition to unveiling association between *APOE* and *HMGCR* polymorphisms, our leading intermediate phenotype of AD susceptibility emphasized the central role of the interaction between brain-derived neurotrophic factor (BDNF) and APOE in influencing HC integrity (Chapter 2, Figs. 1 & 6). The BDNF protein plays a pivotal role in the integrity of hippocampal and neocortical neurons [89]. Post-mortem autopsy revealed that BDNF mRNA expression is decreased in the HC of AD patients as compared to healthy controls [90]. The combined effects of the *APOE* ε4 variant and the *BDNF* Met allele have been shown to precipitate cognitive decline and Aβ deposition in healthy older adults [91]. In contrast, the interaction of the *BDNF* Met allele with the *APOE* ε2 variant is thought to lead to a lesser decline in episodic memory performance when compared to ε3 homozygotes and ε4 carriers [92]. A plausible explanation is that APOE isoforms exert distinct regulatory effects on the maturation and secretion of BDNF. Indeed, human astrocytes lines treated with APOE ε4 secreted negligible amounts of BDNF compared to those treated with the ε2 and ε3 variants [93] The ε4 variant is thought to epigenetically suppresses BDNF mRNA expression by acting on histone acetylation [94]. The impact of APOE on BDNF signaling and its downstream effects on neuronal integrity could be significantly contingent on adverse life events, with enduring repercussions on the HC microstructure. BDNF polymorphism has indeed been shown to influence the relationship between childhood trauma and cognitive performance [95]. This effect was particularly notable in the domains of executive function and verbal fluency, both of which are

especially impacted during the development of AD [95]. Moreover, carriers of the Met allele who experienced childhood abuse exhibited a significant reduction in HC grey matter volume in adulthood [95]. Our cross-generational analysis of AD susceptibility uncovered grey matter alterations in the hippocampus' CA1 and CA3, which are crucial for spatial navigation and episodic memory separation, on the leading intermediate phenotype of AD susceptibility (Chapter 2, Fig. 6). This specific intermediate phenotype has revealed associations between *APOE* and *BDNF* polymorphisms, reinforcing the intricate role of gene-gene interactions in shaping the regulatory effects of the APOE protein on AD risk.

The gene-gene interactions identified in Chapters 1 & 2 appear to either enhance or diminish the isoform-specific effects associated with the APOE protein. Various genetic markers potentially play a role in shaping APOE's homeostatic responses to neurodegeneration and age-related metabolic alterations. This intricate network of connections, in turn, leads to distinctive susceptibility profiles in ε2 and ε4 carriers as compared to ε3 homozygotes. While *APOE* polymorphism appears to play a pivotal role in shaping AD susceptibility in humans, it lacks a counterpart in the animal kingdom, where a singular isoform predominates [96]. The *APOE* ε4 allele, the ancestral haplotype, prevails to this day despite consistently demonstrating an association with AD [24, 97]. The prevalence of the ε3/ε4 haplotype remains substantial, second only to ε3/ε3, ranging from 16 to 41% across different ethnicities [97]. The persistence of the ε3/ε4 haplotype becomes even more perplexing when considering its association with a 2-3 fold increase in AD risk [97]. The enduring yet highly variable distribution of the ε4 allele across diverse environmental niches and human populations poses a significant evolutionary challenge, potentially representing a case of antagonistic pleiotropy. Antagonistic pleiotropy is a concept in evolutionary biology postulating that specific genes or alleles confer disparate fitness consequences throughout distinct life stages of an organism [98]. The emergence of an extended post-reproductive lifespan, a characteristic largely distinctive to humans, has been ascribed to the favorable selection of haplotypes that mitigate age-related cognitive deterioration [99]. The presence of the ε4 allele is associated with increased age-related mortality due to cardiovascular pathology or cognitive decline, as well as heightened all-cause mortality, when compared to individuals without the ε4 allele [100]. Nevertheless, evidence indicates that bearing an ε4 allele could confer a substantial advantage in earlier life stages. Young carriers of the ancestral ε4

haplotype, spanning from childhood to early adulthood and middle age, have shown enhanced memory, executive function, and verbal fluency compared to ε4 non-carriers [101-103]. *APOE* ε4 has also been associated with a reduction in perinatal and infant mortality, along with enhancements in newborn health status, infant cognitive development, and neuronal protection in comparison to *APOE* ε3 [104-106]. Reproductive selection pressure for the ε3 allele might be negligible in a population facing an already higher risk of mortality, independently of *APOE*. The ε4 alleles is indeed relatively more frequent in indigenous populations than in population from European ancestry [107-111]. Hence, the shift over time from the ancestral ε4 allele to the more recent ε3 and ε2 alleles could potentially be attributed to an adaptation to modern lifestyles prompted by a transition away from environments characterized by fluctuating pathogenicity and substantial infection burden. In these earlier conditions, the ε4 allele could have provided enhanced newborn health and decreased perinatal mortality, potentially clarifying its continued prevalence in present times. Genetic diversity within populations is moulded by evolutionary pressures, resulting in the conservation of specific alleles owing to their benefits in particular environmental niches. Isoform-specific effects related to the APOE protein, especially prominent amongst the ε2 and ε4 variants, thus aligns with the concept of evolutionary trade-offs.

Limitations and Future Directions

Brain imaging and classic genetics encounter limitations in their capacity to infer neuronal function, primarily due to their constrained spatial and temporal resolutions. While magnetic resonance imaging (MRI) techniques offer valuable insights into the brain microstructure in-vivo, they lack the granularity to discriminate individual cell types. Distinguishing metabolic processes associated with neurons and glial cells in humans could be pivotal to our understanding of AD pathophysiology. Classical genomic analyses focus on DNA variants which are assumed to be uniform across the body and brain, leading to similar constraints. Emerging evidence indicates significant variability in protein translation across tissues and cell types. APOE isoforms could exert totally distinctive effects on pathways involved with lipid metabolism, inflammation, and Aβ clearance depending on cell types. The tendency to assume homogeneous signals in both brain imaging and classical genetics may impede the recognition of distinct functions among diverse cell populations within the same brain region. A promising avenue for overcoming these limitations is found in the burgeoning field of single-nucleus genetics, which offers unmatched spatial and temporal resolution, cell-type specificity, and precision for studying individual neurons and glial cells. Future research endeavors dedicated to the exploration of cellular-level metabolic alterations, notably within specific subgroups of neurons and glial cells, could shed light on the biochemical underpinnings of neurodegeneration. Such progress, in turn, could act as a catalyst for the formulation of increasingly precise and targeted therapeutic interventions for AD.

Conclusion

Enhancements in quality of life and extended lifespans have brought to light the detrimental impacts of the ε4 allele, fundamentally influencing the trajectory of AD research until the present day. The ε2 allele has emerged as a distinctive player in this evolving narrative, with growing evidence suggesting its potential interplay with various modifiable risk factors associated with AD. Our comprehensive investigation highlights constellations of ε2-related susceptibilities, ranging from neuroticism and social engagement to cardiovascular fitness and cognitive abilities, some of which appear to be passed down from one generation to the next. Sex has emerged as a pivotal determinant in shaping the relationship between these modifiable risk factors and microstructural alterations of the hippocampus and default network. As we continue to uncover the intricate web of influences contributing to AD susceptibility, it becomes increasingly evident that a one-size-fits-all approach is inadequate. Instead, embracing a family-centred approach holds the potential to finely tailor interventions to individual susceptibility profiles, thereby effectively narrowing down the spectrum of heterogeneous predispositions targeted. These insights underscore the necessity for a paradigm shift in our approach to cognitive decline, urging a proactive and personalized exploration of sex-specific protective factors. Shifting research focus toward *APOE* ε2-mediated pathways, which hold potential for targeted therapeutic interventions, opens a promising avenue for advancing precision medicine in AD, thereby harnessing the realm of genetic and biochemical indicators that drive cognitive resilience in older adults.

References

1. Patterson, C., *World alzheimer report 2018.* 2018.
2. Guerchet, M., M. Prince, and M. Prina, *Numbers of people with dementia worldwide: An update to the estimates in the World Alzheimer Report 2015.* 2020.
3. Ferretti, M.T., et al., *Sex differences in Alzheimer disease—the gateway to precision medicine.* Nature Reviews Neurology, 2018. **14**(8): p. 457-469.
4. *2018 Alzheimer's disease facts and figures.* Alzheimer's & Dementia, 2018. **14**(3): p. 367-429.
5. Bamford, S.-M. and T. Walker, *Women and dementia – not forgotten.* Maturitas, 2012. **73**(2): p. 121-126.
6. Proctor, G., *Listening to Older Women with Dementia: Relationships, voices and power.* Disability & Society, 2001. **16**(3): p. 361-376.
7. Pradier, C., et al., *The Mini Mental State Examination at the time of Alzheimer's disease and related disorders diagnosis, according to age, education, gender and place of residence: a cross-sectional study among the French National Alzheimer database.* PloS one, 2014. **9**(8): p. e103630.
8. Mishra, A. and R.D. Brinton, *Inflammation: Bridging Age, Menopause and APOEε4 Genotype to Alzheimer's Disease.* Front Aging Neurosci, 2018. **10**: p. 312.
9. Pike, C.J., *Sex and the development of Alzheimer's disease.* Journal of Neuroscience Research, 2017. **95**(1-2): p. 671-680.
10. Galea, L.A., et al., *Sex, hormones and neurogenesis in the hippocampus: hormonal modulation of neurogenesis and potential functional implications.* Journal of neuroendocrinology, 2013. **25**(11): p. 1039-1061.
11. Mishra, A., et al., *Dynamic Neuroimmune Profile during Mid-life Aging in the Female Brain and Implications for Alzheimer Risk.* Iscience, 2020. **23**(12): p. 101829.
12. Braak, E. and H. Braak, *Alzheimer's disease: transiently developing dendritic changes in pyramidal cells of sector CA1 of the Ammon's horn.* Acta Neuropathologica, 1997. **93**(4): p. 323-325.
13. Barbas, H. and G.J. Blatt, *Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey.* Hippocampus, 1995. **5**(6): p. 511-533.
14. Amaral, D.G., H.E. Scharfman, and P. Lavenex, *The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies)*, in *Progress in Brain Research*, H.E. Scharfman, Editor. 2007, Elsevier. p. 3-790.
15. Cameron, H.A. and E. Gould, *Adult neurogenesis is regulated by adrenal steroids in the dentate gyrus.* Neuroscience, 1994. **61**(2): p. 203-209.
16. Tanapat, P., et al., *Estrogen stimulates a transient increase in the number of new neurons in the dentate gyrus of the adult female rat.* Journal of Neuroscience, 1999. **19**(14): p. 5792-5801.
17. Mazzucco, C.A., et al., *Both estrogen receptor α and estrogen receptor β agonists enhance cell proliferation in the dentate gyrus of adult female rats.* Neuroscience, 2006. **141**(4): p. 1793-1800.
18. Hajszan, T., T.A. Milner, and C. Leranth, *Sex steroids and the dentate gyrus*, in *Progress in Brain Research*, H.E. Scharfman, Editor. 2007, Elsevier. p. 399-816.

19. Brinton, R.D., *The healthy cell bias of estrogen action: mitochondrial bioenergetics and neurological implications.* Trends in Neurosciences, 2008. **31**(10): p. 529-537.
20. Prestwood, K.M., et al., *The effect of different doses of micronized 17β-estradiol on C-reactive protein, interleukin-6, and lipids in older women.* The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 2004. **59**(8): p. M827-M832.
21. Thelen, K.M., et al., *Cholesterol synthesis rate in human hippocampus declines with aging.* Neuroscience Letters, 2006. **403**(1): p. 15-19.
22. Svennerholm, L. and C.-G. Gottfries, *Membrane Lipids, Selectively Diminished in Alzheimer Brains, Suggest Synapse Loss as a Primary Event in Early-Onset Form (Type I) and Demyelination in Late-Onset Form (Type II).* Journal of Neurochemistry, 1994. **62**(3): p. 1039-1047.
23. Eckert, G.P., et al., *Cholesterol Modulates the Membrane- Disordering Effects of Beta-Amyloid Peptides in the Hippocampus: Specific Changes in Alzheimer's Disease.* Dementia and Geriatric Cognitive Disorders, 2000. **11**(4): p. 181-186.
24. Poirier, J., et al., *Apolipoprotein E polymorphism and Alzheimer's disease.* The Lancet, 1993. **342**(8873): p. 697-699.
25. Poirier, J., *Apolipoprotein E, cholesterol transport and synthesis in sporadic Alzheimer's disease.* Neurobiology of aging, 2005. **26**(3): p. 355-361.
26. Fullerton, S.M., et al., *Apolipoprotein E Variation at the Sequence Haplotype Level: Implications for the Origin and Maintenance of a Major Human Polymorphism.* The American Journal of Human Genetics, 2000. **67**(4): p. 881-900.
27. Huebbe, P. and G. Rimbach, *Evolution of human apolipoprotein E (APOE) isoforms: Gene structure, protein function and interaction with dietary factors.* Ageing Research Reviews, 2017. **37**: p. 146-161.
28. Mahley, R.W., *Apolipoprotein E: cholesterol transport protein with expanding role in cell biology.* Science, 1988. **240**(4852): p. 622-630.
29. Kim, Y.J., et al., *Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: a three-year longitudinal study.* Scientific reports, 2017. **7**(1): p. 1-8.
30. Jansen, W.J., et al., *Prevalence of Cerebral Amyloid Pathology in Persons Without Dementia: A Meta-analysis.* JAMA, 2015. **313**(19): p. 1924-1938.
31. Nagy, Z.S., et al., *Influence of the apolipoprotein E genotype on amyloid deposition and neurofibrillary tangle formation in Alzheimer's disease.* Neuroscience, 1995. **69**(3): p. 757-761.
32. Lim, Y.Y., E.C. Mormino, and A.s.D.N. Initiative, *APOE genotype and early β-amyloid accumulation in older adults without dementia.* Neurology, 2017. **89**(10): p. 1028-1034.
33. Strittmatter, W.J., et al., *Binding of human apolipoprotein E to synthetic amyloid beta peptide: isoform-specific effects and implications for late-onset Alzheimer disease.* Proceedings of the National Academy of Sciences, 1993. **90**(17): p. 8098-8102.
34. Seshadri, S., et al., *Lifetime risk of dementia and Alzheimer's disease.* The impact of mortality on risk estimates in the Framingham Study, 1997. **49**(6): p. 1498-1504.
35. Altmann, A., et al., *Sex modifies the APOE-related risk of developing Alzheimer disease.* Annals of Neurology, 2014. **75**(4): p. 563-573.
36. Yin, F., et al., *The perimenopausal aging transition in the female rat brain: decline in bioenergetic systems and synaptic plasticity.* Neurobiology of Aging, 2015. **36**(7): p. 2282-2295.

37. Klosinski, L.P., et al., *White matter lipids as a ketogenic fuel supply in aging female brain: implications for Alzheimer's disease.* EBioMedicine, 2015. **2**(12): p. 1888-1904.
38. Aggleton, J.P., et al., *Complementary Patterns of Direct Amygdala and Hippocampal Projections to the Macaque Prefrontal Cortex.* Cerebral Cortex, 2015. **25**(11): p. 4351-4373.
39. Raichle, M.E., et al., *A default mode of brain function.* Proceedings of the National Academy of Sciences, 2001. **98**(2): p. 676-682.
40. Bartzokis, G., et al., *Apolipoprotein E Genotype and Age-Related Myelin Breakdown in Healthy Individuals: Implications for Cognitive Decline and Dementia.* Archives of General Psychiatry, 2006. **63**(1): p. 63-72.
41. Brickman, A.M., et al., *Testing the white matter retrogenesis hypothesis of cognitive aging.* Neurobiology of Aging, 2012. **33**(8): p. 1699-1715.
42. Reiman, E.M., et al., *Correlations between apolipoprotein E ε4 gene dose and brain-imaging measurements of regional hypometabolism.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(23): p. 8299-8302.
43. Reiman, E.M., et al., *Preclinical Evidence of Alzheimer's Disease in Persons Homozygous for the ε4 Allele for Apolipoprotein E.* New England Journal of Medicine, 1996. **334**(12): p. 752-758.
44. Palmqvist, S., et al., *Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity.* Nature communications, 2017. **8**(1): p. 1-13.
45. Damoiseaux, J.S., et al., *Gender modulates the APOE ε4 effect in healthy older adults: convergent evidence from functional brain connectivity and spinal fluid tau levels.* Journal of Neuroscience, 2012. **32**(24): p. 8254-8262.
46. Calderón-Garcidueñas, L., 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.* Environmental research, 2016. **146**: p. 404-417.
47. Cowan, D.F., *Amyloidosis in the Bottlenose Dolphin, Tursiops truncatus.* Veterinary Pathology, 1995. **32**(3): p. 311-314.
48. Smallwood, J., et al., *The default mode network in cognition: a topographical perspective.* Nature Reviews Neuroscience, 2021. **22**(8): p. 503-513.
49. Lu, H., et al., *Rat brains also have a default mode network.* Proceedings of the National Academy of Sciences, 2012. **109**(10): p. 3979-3984.
50. Rilling, J.K., et al., *A comparison of resting-state brain activity in humans and chimpanzees.* Proceedings of the National Academy of Sciences, 2007. **104**(43): p. 17146-17151.
51. Buckner, R.L., J.R. Andrews-Hanna, and D.L. Schacter, *The brain's default network: Anatomy, function, and relevance to disease*, in *The year in cognitive neuroscience 2008.* 2008, Blackwell Publishing: Malden. p. 1-38.
52. Addis, D.R., A.T. Wong, and D.L. Schacter, *Remembering the past and imagining the future: Common and distinct neural substrates during event construction and elaboration.* Neuropsychologia, 2007. **45**(7): p. 1363-1377.
53. Saxe, R., *The right temporo-parietal junction: a specific brain region for thinking about thoughts.* Handbook of theory of mind, 2010: p. 1-35.

54. Kernbach, J.M., et al., *Subspecialization within default mode nodes characterized in 10,000 UK Biobank participants.* Proceedings of the National Academy of Sciences, 2018. **115**(48): p. 12295-12300.
55. O'keefe, J. and L. Nadel, *The hippocampus as a cognitive map*. 1978: Oxford university press.
56. Bubb, E.J., C. Metzler-Baddeley, and J.P. Aggleton, *The cingulum bundle: anatomy, function, and dysfunction.* Neuroscience & Biobehavioral Reviews, 2018. **92**: p. 104-127.
57. Swanson, L.W. and W.M. Cowan, *Hippocampo-Hypothalamic Connections: Origin in Subicular Cortex, Not Ammon's Horn.* Science, 1975. **189**(4199): p. 303-304.
58. Oyola, M.G. and R.J. Handa, *Hypothalamic–pituitary–adrenal and hypothalamic–pituitary–gonadal axes: sex differences in regulation of stress responsivity.* stress, 2017. **20**(5): p. 476-494.
59. Hamilton, J.P., et al., *Default-Mode and Task-Positive Network Activity in Major Depressive Disorder: Implications for Adaptive and Maladaptive Rumination.* Biological Psychiatry, 2011. **70**(4): p. 327-333.
60. Zhou, H.-X., et al., *Rumination and the default mode network: Meta-analysis of brain imaging studies and implications for depression.* NeuroImage, 2020. **206**: p. 116287.
61. Perkins, A.M., et al., *Thinking too much: Self-generated thought as the engine of neuroticism.* Trends in cognitive sciences, 2015. **19**(9): p. 492-498.
62. Zajner, C., R.N. Spreng, and D. Bzdok, *Loneliness is linked to specific subregional alterations in hippocampus-default network covariation.* Journal of Neurophysiology, 2021. **126**(6): p. 2138-2157.
63. Shafighi, K., et al., *Social isolation is linked to classical risk factors of Alzheimer's disease-related dementias.* bioRxiv, 2021.
64. Zhou, Z., P. Wang, and Y. Fang, *Loneliness and the risk of dementia among older Chinese adults: gender differences.* Aging & Mental Health, 2018. **22**(4): p. 519-525.
65. Calderón-Garcidueñas, L., et al., *Neuroinflammation, Hyperphosphorylated Tau, Diffuse Amyloid Plaques, and Down-Regulation of the Cellular Prion Protein in Air Pollution Exposed Children and Young Adults.* Journal of Alzheimer's Disease, 2012. **28**: p. 93-107.
66. Calderón-Garcidueñas, L., et al., *Decreases in short term memory, IQ, and altered brain metabolic ratios in urban apolipoprotein ε4 children exposed to air pollution.* Journal of Alzheimer's Disease, 2015. **45**(3): p. 757-770.
67. Podewils, L.J., et al., *Physical activity, APOE genotype, and dementia risk: findings from the Cardiovascular Health Cognition Study.* American journal of epidemiology, 2005. **161**(7): p. 639-651.
68. Levy, B.R., et al., *When Culture Influences Genes: Positive Age Beliefs Amplify the Cognitive-Aging Benefit of APOE ε2.* The Journals of Gerontology: Series B, 2020. **75**(8): p. e198-e203.
69. Buckner, R.L., et al., *Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory.* Journal of neuroscience, 2005. **25**(34): p. 7709-7717.
70. Sperling, R.A., et al., *Amyloid deposition is associated with impaired default network function in older persons without dementia.* Neuron, 2009. **63**(2): p. 178-188.

71. Tremblay-Mercier, J., et al., *Open science datasets from PREVENT-AD, a longitudinal cohort of pre-symptomatic Alzheimer's disease.* NeuroImage: Clinical, 2021. **31**: p. 102733.
72. Helmer, M., et al., *On stability of canonical correlation analysis and partial least squares with application to brain-behavior associations.* BioRxiv, 2021: p. 2020.08. 25.265546.
73. Wang, H.-T., et al., *Finding the needle in a high-dimensional haystack: Canonical correlation analysis for neuroscientists.* NeuroImage, 2020. **216**: p. 116745.
74. Shinohara, M., et al., *APOE2 eases cognitive decline during aging: clinical and preclinical evaluations.* Annals of neurology, 2016. **79**(5): p. 758-774.
75. Riddell, D.R., et al., *Impact of Apolipoprotein E (ApoE) Polymorphism on Brain ApoE Levels.* The Journal of Neuroscience, 2008. **28**(45): p. 11445-11453.
76. Ulrich, J.D., et al., *In vivo measurement of apolipoprotein E from the brain interstitial fluid using microdialysis.* Molecular Neurodegeneration, 2013. **8**(1): p. 13.
77. Beffert, U., 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 Research, 1999. **843**(1): p. 87-94.
78. Krasemann, S., et al., *The TREM2-APOE Pathway Drives the Transcriptional Phenotype of Dysfunctional Microglia in Neurodegenerative Diseases.* Immunity, 2017. **47**(3): p. 566-581.e9.
79. Stone, D.J., et al., *Astrocytes and Microglia Respond to Estrogen with Increased apoE mRNAin Vivoandin Vitro.* Experimental Neurology, 1997. **143**(2): p. 313-318.
80. Lambert, J., N. Coyle, and C. Lendon, *The allelic modulation of apolipoprotein E expression by oestrogen: potential relevance for Alzheimer's disease.* Journal of medical genetics, 2004. **41**(2): p. 104-112.
81. Mahley, R.W., *Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders.* Journal of Molecular Medicine, 2016. **94**(7): p. 739-746.
82. Podcasy, J.L. and C.N. Epperson, *Considering sex and gender in Alzheimer disease and other dementias.* Dialogues in clinical neuroscience, 2016. **18**(4): p. 437-446.
83. Goldstein, J.L. and M.S. Brown, *Regulation of the mevalonate pathway.* Nature, 1990. **343**(6257): p. 425-430.
84. Wall, C.T.J., et al., *Mitochondrial respiratory chain dysfunction alters ER sterol sensing and mevalonate pathway activity.* Journal of Biological Chemistry, 2022. **298**(3).
85. Karlic, H. and F. Varga, *Mevalonate Pathway*, in *Encyclopedia of Cancer (Third Edition)*, P. Boffetta and P. Hainaut, Editors. 2019, Academic Press: Oxford. p. 445-457.
86. Leduc, V., et al., *HMGCR is a genetic modifier for risk, age of onset and MCI conversion to Alzheimer's disease in a three cohorts study.* Molecular Psychiatry, 2015. **20**(7): p. 867-873.
87. Fester, L., et al., *Cholesterol-promoted synaptogenesis requires the conversion of cholesterol to estradiol in the hippocampus.* Hippocampus, 2009. **19**(8): p. 692-705.
88. Belloy, M.E., V. Napolioni, and M.D. Greicius, *A quarter century of APOE and Alzheimer's disease: progress to date and the path forward.* Neuron, 2019. **101**(5): p. 820-838.
89. Murer, M.G., Q. Yan, and R. Raisman-Vozari, *Brain-derived neurotrophic factor in the control human brain, and in Alzheimer's disease and Parkinson's disease.* Progress in neurobiology, 2001. **63**(1): p. 71-124.

90. Phillips, H.S., et al., *BDNF mRNA is decreased in the hippocampus of individuals with Alzheimer's disease.* Neuron, 1991. **7**(5): p. 695-702.
91. Lim, Y.Y., et al., *APOE and BDNF polymorphisms moderate amyloid β-related cognitive decline in preclinical Alzheimer's disease.* Molecular Psychiatry, 2015. **20**(11): p. 1322-1328.
92. Ward, D.D., et al., *APOE and BDNF Val66Met polymorphisms combine to influence episodic memory function in older adults.* Behavioural Brain Research, 2014. **271**: p. 309-315.
93. Sen, A., T.J. Nelson, and D.L. Alkon, *ApoE isoforms differentially regulates cleavage and secretion of BDNF.* Molecular Brain, 2017. **10**(1): p. 19.
94. Abhik, S., J.N. Thomas, and L.A. Daniel, *ApoE4 and Aβ Oligomers Reduce BDNF Expression via HDAC Nuclear Translocation.* The Journal of Neuroscience, 2015. **35**(19): p. 7538.
95. Aas, M., et al., *BDNF val66met modulates the association between childhood trauma, cognitive and brain abnormalities in psychoses.* Progress in Neuro-Psychopharmacology and Biological Psychiatry, 2013. **46**: p. 181-188.
96. Zannis, V., et al., *Plasma and hepatic apoE isoproteins of nonhuman primates. Differences in apoE among humans, apes, and New and Old World monkeys.* Journal of lipid research, 1985. **26**(12): p. 1421-1430.
97. Farrer, L.A., 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): p. 1349-1356.
98. Tuminello, E.R. and S.D. Han, *The apolipoprotein e antagonistic pleiotropy hypothesis: review and recommendations.* International journal of Alzheimer's disease, 2011. **2011**.
99. Schwarz, F., et al., *Human-specific derived alleles of CD33 and other genes protect against postreproductive cognitive decline.* Proceedings of the National Academy of Sciences, 2016. **113**(1): p. 74-79.
100. Ewbank, D.C., *The APOE gene and differences in life expectancy in Europe.* The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 2004. **59**(1): p. B16-B20.
101. Mondadori, C.R.A., et al., *Better Memory and Neural Efficiency in Young Apolipoprotein E ε4 Carriers.* Cerebral Cortex, 2006. **17**(8): p. 1934-1947.
102. Marchant, N.L., et al., *Positive Effects of Cholinergic Stimulation Favor Young APOE ε4 Carriers.* Neuropsychopharmacology, 2010. **35**(5): p. 1090-1096.
103. Alexander, D.M., et al., *The contribution of apolipoprotein E alleles on cognitive performance and dynamic neural activity over six decades.* Biological Psychology, 2007. **75**(3): p. 229-238.
104. Wright, R.O., et al., *Apolipoprotein E genotype predicts 24-month bayley scales infant development score.* Pediatric research, 2003. **54**(6): p. 819-825.
105. Becher, J.-C., et al., *Apolipoprotein E e4 and its prevalence in early childhood death due to sudden infant death syndrome or to recognised causes.* Early Human Development, 2008. **84**(8): p. 549-554.
106. Becher, J.-C., et al., *The distribution of apolipoprotein E alleles in Scottish perinatal deaths.* Journal of medical genetics, 2006. **43**(5): p. 414-418.
107. Gerdes, L.U., et al., *The apolipoprotein E polymorphism in Greenland Inuit in its global perspective.* Human Genetics, 1996. **98**: p. 546-550.

108. Lehtinen, S., et al., *Differences in genetic variation of apolipoprotein E in Lapps and Finns.* Atherosclerosis, 1994. **109**(1-2): p. 263.
109. Kamboh, M., S. Serjeantson, and R. Ferrell, *Genetic studies of human apolipoproteins. XVIII. Apolipoprotein polymorphisms in Australian Aborigines.* Human biology, 1991: p. 179-186.
110. Kamboh, M., K.K. Bhatia, and R. Ferrell, *Genetic studies of human apolipoproteins: XII. Population genetics of apolipoproteins in Papua New Guinea.* American Journal of Human Biology, 1990. **2**(1): p. 17-23.
111. Gajra, B., et al., *Effect of apolipoprotein E variants on plasma lipids and apolipoproteins in the Orang Asli ('Aborigines') of Malaysia.* Human heredity, 1994. **44**(4): p. 209-213.