Genetic and Cardiometabolic Contributions to Cognitive, Structural Brain and Biomarker Phenotypes

Rachana Tank

BA (Hons), MSc

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Institute of Health and Wellbeing College of Medical, Veterinary and Life Sciences

University of Glasgow

January 2022

Abstract

The number of individuals experiencing abnormal cognitive ageing is rapidly increasing, which can only in part be explained by an ageing population. Prevention of considerable cognitive decline is complicated by its heterogeneity and numerous risk factors. The most significant contributions to brain health and cognitive decline outside of age are higher genetic risk and poor cardiometabolic health. There are gaps in the literature and understanding regarding the extent to which common genetic or cardiometabolic conditions contribute to and interact with one another to influence brain health. Therefore, the overall aim of this PhD project is to explore genetic and cardiometabolic risks in relation to structural brain MRI measures, cognitive assessments, and blood biomarkers.

This thesis used large-scale secondary data from the UK Biobank in which several analyses investigating associations between cardiometabolic conditions, genetic risks, cognition, and brain MRI data are the largest to date. The use of the UK Biobank cohort also allowed for controlling of confounders that have not been considered or accounted for in previous studies. The primary focus of this thesis was on genetic and cardiometabolic contributions to cognition and brain MRI. The main objectives were: (1) Contribute to the understanding of cardiovascular to brain health and (2) Determine the role of genetic risk factors on the brain and physical health in healthy adults.

When examining multimorbidity, there were no clear trends between cardiometabolic groups and brain MRI metrics. However, this may have been due to a healthy selection bias in which those with multimorbidity were healthy enough for MRI assessments. When examining genetically elevated risk of cardiovascular disease (CVD) indexed by lipoprotein A (), we found associations with mean diffusivity and fractional anisotropy, suggesting a potential role of LpA in brain ageing. However, we found discrepancies between genetically elevated LpA and blood LpA, which should be further investigated.

When calculating genetic risk scores for Alzheimer's disease (AD) in healthy midlife adults, we found evidence for potential early ageing pathology within subfields of the hippocampus prior to significant cognitive impairments. We also found that this elevated genetic risk for AD was associated with elevated cystatin c. Elevated genetic risk of AD also showed significant sex differences in biomarker analyses. Creatinine and oestradiol were significantly associated with an elevated risk of Alzheimer's in women but not men. These findings support routine stratification in exploratory research.

This thesis emphasises the importance of epidemiological research that considers cardiometabolic, lifestyle and genetic risk factors together in the context of cognitive health. There is scope to build on this work in omics and cohort studies.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my principal supervisor, Donald. I would never have begun (let alone completed) this thesis without you. It has been an absolute pleasure to learn from you over these years. Thank you, Donald, for providing me with opportunities I did not think I had.

Thank you, Kristin for being a fantastic secondary supervisor, you have challenged me to think more deeply with your feedback. Thank you to my third supervisor, Jonathan for reading my work and the consistent optimism.

I would like to thank the UK Biobank research group - being part of this group has been a valuable environment to learn good research and work ethic. Thank you, Joey, in particular for your patience, willingness to help, and in house pipelines that made so much of my thesis possible.

To all members of the PhQueens past and present, gaining a group of hilarious, smart, and kind friends has been a highlight of this degree. Kirstin and Mariana – doing a PhD during a pandemic has not been easy but I am grateful to have never been alone - thank you both for being the loveliest people to share a home with. Thank you, Keira for the much-needed lockdown walks.

Matthew, thank you for the company, conversations and adventurous weekends.

Above all, I thank my parents and my brother for everything we have and whose strength and resilience I have always admired. For that and much more I will always be grateful. This is dedicated to you.

I would also like to thank the UK Biobank participants who dedicated their time to make this data possible. The UK Biobank will be a valuable resource for decades to come, and I am grateful to have had the opportunity to work with it.

Author's Declaration

I declare that I am the sole author of this thesis, except where the assistance of others has been acknowledged. The work in this thesis has not been submitted in any form for another degree or professional qualification.

Rachana Tank
January 2022

Definitions/Abbreviations

Aß	Amyloid beta (plaques)
AD	Alzheimer's Disease
APOE	Apolipoprotein-e
BAG	Brain Age Gap
BMI	Body Mass Index
CAD	Coronary Artery Disease
CMD	Cardiometabolic Disease
CNS	Central Nervous System
CRP	C-Reactive Protein
CSF	Cerebrospinal Fluid
CVD	Cardiovascular Disease
CysC	Cystatin C
EAF	Effect Allele Frequency
FA	Fractional Anisotropy
GM	Grey Matter
GWAS	Genome-wide Association Study
HDL	High-Density Lipoprotein
HT	Hypertension
HWE	Hardy-Weinberg Equilibrium
ICC	Intraclass Correlation
LD	Linkage Disequilibrium
LDL	Low-Density Lipoprotein
LpA	Lipoprotein A
MAF	Minor Allele Frequency
MCI	Mild Cognitive Impairment
MD	Mean Diffusivity
MR	Mendelian Randomization
MRI	Magnetic Resonance Imaging
PC	Principal Component
PGR	Polygenic Risk
RT	Reaction Time
SNP	Single Nucleotide Polymorphism
SD	Standard Deviation
TMT	Trail Making Task
T2D	Type 2 Diabetes
WM	White Matter
WMH	White Matter Hyperintensities
UK	United Kingdom

Chapter 1: Introduction

1.1 Ageing population

Population ageing is defined as an increase in median age in a population and has been recognised as one of four current global demographic trends alongside population growth, international migration, and urbanisation (Leeson, 2018; United Nations, 2019). The global number of individuals over 65 is expected to double to 1.5 billion by 2050, representing major achievements for medicine, social development, and public health. Increasing longevity and declining fertility are the most significant contributing factors to population ageing and determinants of continuing population ageing. An ageing population will require adapting patterns of government support, i.e., mental health services, adequate housing, efforts to reduce social isolation, health literacy and income support. Governments must meet these needs while ensuring that existing health and economic inequalities do not grow, as those in the lowest socio-economic groups are the most likely to experience multiple health problems and find it the hardest to access health services.

Population ageing will result in a greater prevalence of both age-related conditions and individuals living longer with existing conditions with existing conditions, as illustrated in Figure 1.1. One such example is with dementia: 1.7 million individuals are projected to live with dementia in the UK by 2051, a substantial increase on 850,000 individuals in 2015, and the number of individuals over 65 living with coronary heart disease, dementia, diabetes, hypertension and high cholesterol is also projected to rise (Ageing.ox.ac.uk., 2021).

The implications of this demographic shift call for further knowledge about age-related conditions and their risk factors in which stratified medicine and genomics is one such research approach that allows for the discovery of disease pathways. This approach will be useful in targeting or preventing the leading health conditions likely to affect the quality of life in an ageing population, including brain health, cognitive health, and addressing key risk conditions, comorbidities, and mechanisms mediating brain ageing.

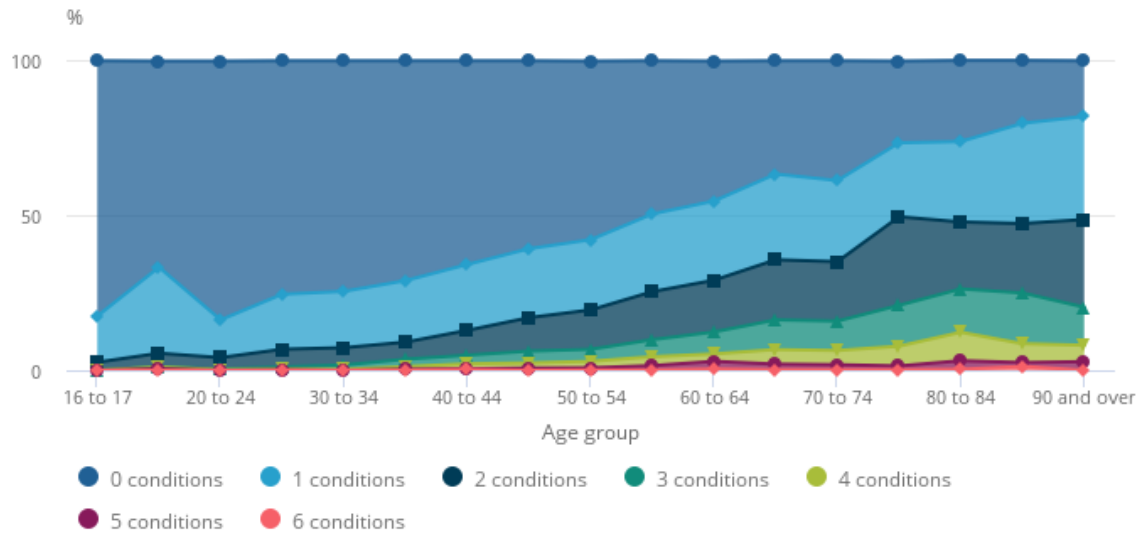

Figure 1.1 Number of grouped health conditions by age
Source: Health Survey for England, NHS Digital, 2016, England

1.1.1 Cognitive health in ageing

Age-related changes to cognition are expected over the lifetime but steep declines are not always typical or expected features of ageing. As the proportion of older adults increases it is important to consider who may be at risk of steep cognitive decline or abnormal cognitive decline. Structural brain and cognitive changes of healthy ageing have been well documented in the literature. Cognitive abilities such as crystallised intelligence typically remain resilient during normal ageing (Salthouse, 2010; 2011). Crystalised abilities are skills, abilities and knowledge that are learned – they accumulate over a lifetime and are based on experiences. Examples of crystallised abilities include reading comprehension, facts or general information. In comparison, fluid intelligence or abilities are understood to decline in normal ageing. This is considered the innate ability to process novel information and problem solve that depends minimally on prior learning or formal education. The cognitive domains considered to be fluid include executive function, processing speed, memory, and psychomotor ability and are observed to decline after the third decade of life. Fluid intelligence is of interest in abnormal ageing as steeper declines are seen in conditions such as dementia. Memory is another cognitive

ability sensitive to both normal and abnormal ageing. Recognition (retrieving memory after a cue), temporal (correct recall of time or sequence of past events) and procedural (how to do things) memory tend to remain intact with normal ageing. In contrast, free recall (retrieval of information with no cue), source memory (knowing the source of the memory) and prospective memory (remembering to perform intended actions in the future) tend to decline as part of normal cognitive ageing (Harada et al., 2013).

When individuals age abnormally, cognitive abilities that typically remain intact may also begin to deteriorate. Normal age-related declines tend to be subtle and affect the speed of thinking and attentional control. However, in abnormal aging, cognitive decline is more severe and may include other abilities, such as rapid forgetting or difficulties navigating, solving common problems or expressing oneself verbally. It is important to note, however, slowed speed of information processing accounts for a large proportion of the age-related decline in all cognitive domains in which the spectrum of cognitive decline ranges from normal cognitive ageing to the dementias.

Alongside cognitive deficits seen in abnormal ageing, brain tissue damage (as measured with structural magnetic resonance imaging, MRI) and changes in neural activity (as measured with functional MRI) are often reflected. However, the literature has consistently shown structural brain changes alone aren't able to provide a full picture at predicting those who will decline. Cognitive reserve may help explain why some individuals can withstand age-related and pathological brain changes vital in maintaining their cognitive functioning. Cognitive reserve is a concept that refers to individual difference in cognitive functioning when changes occur in structural and functional brain mechanisms (Stern et al., 2020). Sociobehavioural proxies have been found to contribute to cognitive reserve in the form of nutrition, sleep and exercise (Radanovic, 2020). A cognitive footprint refers to the concept that over the lifetime an individual may have experiences that impair or enhance cognition which may indicate their cognitive health. These behaviours play a role in brain reserve and maintenance and have been referred to

as cognitive footprint (Rossor & Knapp, 2015). It is becoming clear that these lifestyle factors must be considered alongside cognition in order to understand which individuals are most at risk of deteriorating cognition. Such variables can be considered in cohort studies like UK Biobank where there are sociodemographic data available.

1.2 Vascular risk factors for brain and cognitive ageing

Risk factors and determinants are defined as variables associated with an increased risk of disease or decline, for which there is a large body of evidence demonstrating that the structure and integrity of the ageing brain are closely related to physical health. For decades, the relationship between vascular health and brain health has been discussed, with cerebrovascular alterations being a common aetiology in abnormal cognitive decline in ageing (Knopman et al., 2005). The most common presentation of impaired cerebrovascular health is stroke, with high blood pressure being the most significant risk factor. However, cerebrovascular health refers to various conditions that affect the blood vessels and cerebral circulation of the brain and decline in cerebrovascular health can occur due to cardiovascular and cardiometabolic conditions or associated risk factors. Cardiometabolic disease describes a spectrum of cardiovascular and metabolic related conditions which encompasses cardiovascular diseases, but also includes conditions such as insulin resistance and diabetes. Cardiovascular disease can be considered one of many cardiometabolic conditions where metabolic specific dysfunctions are characterised by insulin resistance, impaired glucose tolerance, dyslipidemia, and central adiposity. Cardiovascular and cardiometabolic conditions both include hypertension and coronary disease, whereas diabetes and ischemic stroke are cardiometabolic and cardiovascular conditions respectively. Metabolic and cardiovascular disease risk factors are well understood to increase the risk of decline in cerebrovascular health (Hegele et al., 2010; Marchant et al., 2012; Hooghiemstra et al., 2017; Gupta et al., 2018; Arnoldussen et al., 2019). This thesis does not use the terms

interchangeably, but uses both depending on the most relevant term, or the one that is commonly referred to in the area of research or papers.

Impairment of cerebrovascular health causes damage to both cerebral tissue (grey matter) and white matter, often presenting on MRI scans as white matter hyperintensities (WMH). WMH are visible as areas of increased brightness in MRI T2 weighted images, as MRI is sensitive to small changes in water content within the brain. The initial discovery of WMHs was made in 1987 by Hachinski et al. in which they described WMH as patchy low attenuation in the periventricular and deep white matter of the brain. However, WMH can also present as small subcortical infarcts, lacunes (fluid-filled cavity), microbleeds, atrophy, and also as hyperintense periventricular spaces or lesions (Wardlaw et al., 2015). Presentation of WMH on MRI scans are the hallmark of cerebrovascular impairments and have also consistently been associated with cognitive impairments (Dufouil et al., 2001; Wang et al., 2015; Wardlaw et al., 2015; Bangen et al., 2018). These abnormal cognitive changes in ageing found with vascular risk factors are thought to be due to changes to the myelination of neurons, number of synaptic connections, availability of neurotransmitters or cerebral perfusion. However, WMH are found in a significant proportion of cognitively normal elderly populations and may initially present with no neuropsychological symptoms. They can also be seen in autoimmune diseases and psychiatric illnesses (Wardlaw et al., 2015).

Studies examining how prevalent conditions such as cardiovascular disease (CVD) or cardiometabolic disease (CMD) affect structural brain integrity in older age have provided insights into mechanisms through which cognitive decline occurs. As a result, there is substantial evidence documenting the associations between vascular risk factors and cognitive consequences. Lyall et al. (2017) looked at differences between single and multiple cardiometabolic conditions and cognitive abilities in the UK Biobank and found an additive association between cardiometabolic conditions and cognitive functions, specifically processing and reasoning domains. Cohen et al. (2009) found atherosclerotic burden within blood vessels was associated with lower cognitive function and higher WMH volume. Li et al., (2011)

found that vascular risk factors were also valuable in predicting those with mild cognitive impairment (MCI) who would go on to develop Alzheimer's dementia (AD). Another study by Espana-Irla et al. (2021) found in healthy middle-aged adults (40-65 years) that cardiovascular health measured by VO2 (maximal oxygen consumption) and a cardiovascular risk assessment were associated with several clinical neurocognitive assessments, including visuospatial reasoning (b=-0.046, p=0.002), processing speed (b=-0.115, p=<0.001) and memory (b=-0.120, p=<0.001). Other landmark studies examining vascular health and cognition, or brain structure include Murray et al., (2005), Dickstein et al. (2010), The ARIC Study (Knopman et al., 2011) and Cox et al. (2019).

Many studies in this area highlight the importance of considering lifestyle factors when studying cardiovascular, brain and cognitive associations. Younger and healthier middle-aged adults may be of interest to study as modifiable factors (such as tobacco and alcohol use, or inactivity) contribute to cognitive and brain ageing.

1.2.1 Anatomical and MRI substrates of vascular risk factors

Some studied examining ageing phenotypes have investigated differences between chronological and brain predicted age with the goal of elucidating biological mechanisms of neurodegeneration. This is referred to as the brain age gap (BAG), with larger BAGs being reported in individuals with psychiatric or neurological disorders. Kolbeinsson et al. (2020) looked at 21,382 individuals from UK Biobank and calculated brain ages. Six brain regions were found to contribute to accuracy of age prediction, including left cerebellar lobules I-IV, left crus and vermis, right hippocampus, left amygdala and left insular cortex. When subsequently examining these brain differences using ICD codes via a phenome-wide association study, 24 variables were associated with brain age differences including cardiovascular and metabolic diseases, their risk factors, cognitive functioning, and physical strength. These analyses showed associations between larger brain age gaps and vascular disease risk factors in healthy individuals from the UK Biobank cohort. Similarly,

Beck et al., (2020) looked at several cardiometabolic risk factors over time, and their respective associations with BAGs. The study found that levels of phosphate, blood pressure, smoking, pulse rate and levels of c-reactive protein (CRP) were associated with older-appearing brains. These results further support that cardiometabolic and vascular risk factors are associated with brain ageing.

There is now a need for a closer understanding of these cardiometabolic conditions' relationship to the brain, as many affect neurodegeneration in vastly different ways. Lamar et al. (2020) carried out two separate literature reviews that hypothesised that a core group of typical brain structural alterations exists between CVD risk factors and Alzheimer's dementia. The study found evidence to support their hypothesis in which 23 regions were commonly associated with both CVD risk factors and Alzheimer's dementia. Friedman et al. (2014) concluded similar findings from 77 studies in individuals without cardiovascular conditions but the presence of CVD risk factors including hypertension, diabetes mellitus, obesity, hyperlipidemia, and smoking. They found that all risk factors were independently associated with brain imaging changes, including whole brain volume reductions, grey matter volume reductions and white matter changes, including the presence of WMH. The findings of a study by Cox et al. (2019) support the above observations. They studied both individual and aggregate measures of vascular risk factors in the UK Biobank including smoking, hypertension, pulse pressure, diabetes, hypercholesterolemia, BMI and waist-hip ratio. They found that the composite vascular risk factor burden score was associated with multiple brain MRI hallmarks of dementia, including WMH volumes and poorer white matter microstructure, with the strongest effects found in the frontal, anterior lateral and medial temporal lobes. These results further support that some brain structures can indicate or reflect "unhealthy" brain ageing, with some regions more affected than others.

Taken together, the current literature provides evidence that cardiometabolic risk factors and comorbidities contribute to abnormal age-related brain changes, highlighting the need to understand how common comorbid conditions interact with and contribute to one another leading to cognitive disability. Particularly as there is

evidence that presence of one condition often indicates an increased risk of metabolically linked multimorbidities. Neuroimaging studies that have investigated commonalities across cardiometabolic conditions and risk factors provide a comprehensive understanding of common mid-life modifiable risk factors and their mechanistic role in the aetiology of ageing relevant phenotypes. Several cohort studies have been dedicated to investigating conditions or specific features of cardiometabolic health concerning brain and cognition. These include the SPRINT MIND (Systolic Blood Pressure Intervention Trial – Memory and cognition in Decreased Hypertension), the Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER) (Kivipelto et al., 2013), the Framingham Heart study (Elena Petrea et al., 2020), and the Atherosclerosis Risk in Communities Study (Wu et al., 2019).

1.3 Genetic risk factors for brain and cognitive ageing

In the context of genetic research, understanding the role of one gene in one condition has historically been considered a major success. However, at present, genetic studies looking at factors influencing ageing have developed rapidly in the past two decades due to emerging technologies such as next-generation sequencing, allowing assays that accommodate over 1 million single nucleotide polymorphisms (SNPs). This has allowed genetics research to study population genetic variation at the level of the genome. It has also allowed for sequencing of the human genome to capture genetic variation in European, African and Asian populations (HapMap) and produce SNP databases such as dbSNP. These efforts have been used to discover the genetic basis of complex diseases and phenotypic traits, particularly phenotypes that are difficult to study such as psychiatric conditions. Biobank datasets with large numbers of variables available for large sample sizes allow for studies in which hypothesis free approaches can be taken. Such analyses have resulted in discovery of age-related phenotypes. One example of such phenotype discovery from a large cohort study is with the UK Biobank imaging subsample where Elliot et al., (2018) examined the genetic architecture of the UK Biobank's image derived phenotypes to understand the genetic architecture of

brain structure. Elliot et al., (2018) carried out GWAS on 3,144 functional and structural brain imaging phenotypes. They found that genes linked to brain development and plasticity tended to be related to mental health disorders, whereas iron-related proteins tended to be related to neurodegenerative disorders, such as Alzheimer's disease. Such studies have allowed for the genetic correlates of disease markers to be identified and contribute to the understanding of underlying genetic architecture of the brain and cognition. They also found that genes linked to brain development and plasticity tended to be related to mental health disorders, whereas iron-related proteins tended to be related to neurodegenerative disorders, such as Alzheimer's disease. Such studies have allowed for the genetic correlates of disease markers to be understood and identified. Additionally, several landmark genome-wide association studies (GWAS) have been published in recent years that have implicated many loci, mechanisms and epigenetics, including GWAS for Alzheimer's dementia (AD) (Kunkle et al., 2019), cognition (Trampush et al., 2017), cerebrovascular health and stroke (Hegele & Dichgans, 2010). The most extensive genetic study to date of cognitive functioning has been published including over 300,000 individuals, identifying over 100 relevant loci, including the implication of cardiometabolic traits such as hypertension (Davies et al., 2018).

The clinical landscape for risk assessment of cognitive ageing or cardiometabolic conditions does not currently consider who may be at a higher genetic risk, however there is evidence showing stratification of individuals by genetic risk can be clinically valuable. A multifactorial risk assessment considering genetics will likely be implemented as we learn more about how to differentiate individuals by genetic traits. However, it will not be possible to implement such tools clinically until we better understand the ways in which single variants influence gene networks and lead to a phenotype. Variants do not work in isolation, but thousands interact and influence each other in large networks; Ronald Fisher (1918) introduced the basis for this idea, suggesting that continuous variation in phenotypes could be the result of Mendelian inheritance. Based on this model, it was suggested most genes have an infinitely small effect on the phenotype of interest and the concept has since been built on by research such as Boyle et al. (2017, p 1), who stated, "most heritability

can be explained by effects on genes outside core pathways". It is now well understood that individual differences in predispositions to complex traits, whether physiological or psychological, are influenced to a greater or lesser degree by genetic factors which tend to be polygenic. One method where a large number of variants can be considered together is by polygenic risk (PGR) scoring. PGR scoring calculates an individual's genetic risk of a trait or disease, such as cardiovascular conditions or ageing phenotypes. PGR scores have been developed for specific subgroups of conditions, including coronary artery disease CAD (Khera et al., 2017), hypertension (Krogager et al., 2018) stroke (Hachiya et al., 2020) and atrial fibrillation (Muse et al., 2018). The Polygenic Score Catalog is an open database of published PGR scores including information about variants, alleles, weights, and relevant GWAS or metadata used. It currently contains around 38 PGR score models for CVD in which the number of contributing SNPs ranges from 27 to 6.5 million.

PGR scoring methods have evolved since the first scores were calculated and now tend to be weighted according to effect sizes and are based on genome-wide SNP data which allows for the integration of more sites of DNA variation. Different statistical approaches can be used to develop a PGR score, in which models are then rigorously validated for their value in predicting a disease state or trait. Weighted scores can be calculated through the use of software that can account for linkage disequilibrium, one of which is LDpred (Vilhjálmsson et al., 2015). However, PGR scoring is not standardised, and many factors can still affect the scores, including GWAS data used, different statistical approaches to the selection and weighting of relevant SNPs, and the process of interpreting the output can also influence findings. This has implications in the subsequent application of polygenic prediction models. However, there remains growing evidence of the clinical utility of stratifying individuals by both clinical and polygenic risk to improve individual risk assessment. Incorporating PGR scores has been shown to perform better than current clinical risk stratification tools and offers more significant opportunities for earlier intervention (Slunecka et al., 2021). This is of particular interest with disease prevention as genetic risk is present at birth, and as more sensitive and accurate PGR tools are developed, it may help prevent younger individuals from

developing classical risk factors. This preventative approach can lead to fewer individuals with clinical events, healthier ageing and less cognitive decline.

One area where PGR scoring methods have shown to work well is within cardiovascular health to index cardiovascular risk. A landmark study by Khera et al. (2017) was notable as they used a PGR score for CAD using GWAS summary statistics (Nickpay et al., 2015). They found when using this PGR score in UK Biobank that individuals in the top 1% of genetic risk for CAD had a 5-fold risk of developing CAD. This was similar to findings by Abraham et al., (2019) also showing that those in the top 1% of polygenic risk were at 3-fold risk of a stroke. When this PGR score was examined in other prospective cohorts, including the Women's Genome Health Study, adherence to healthy lifestyle was associated with a 50% reduction in disease risk in particular for those in the highest quintile of polygenic risk of stroke suggesting that risk can be modified via lifestyle strategies in those most at risk. It also highlights the importance and role of lifestyle factors when considering genetic risk, as it is not the complete picture.

There is also growing evidence that stratifying by risk can not only be used in joint decision making in treatment options such as medication, but also provides greater opportunity to further stratify individuals according to other risks, enabling the clarification of relative roles of risk factors. For example, Riveros-Mckay et al. (2021) found that polygenic risk showed strong predictive power for CAD events, which further improved when individuals were stratified into age versus sex subgroups. When studying longevity phenotypes, Melzer et al. (2020) found that when stratifying those in the top 10% of parental survival age, a GWAS showed *APOE*, *CHRNA3*, *LPA* and *CDKN2B-AS1* were significantly associated, which are all also implicated in cardiometabolic risk. This implies there is some degree of genetic overlap between extreme longevity and cardiometabolic health, which is of potential interest for assessing risk for cognitive ageing and brain health.

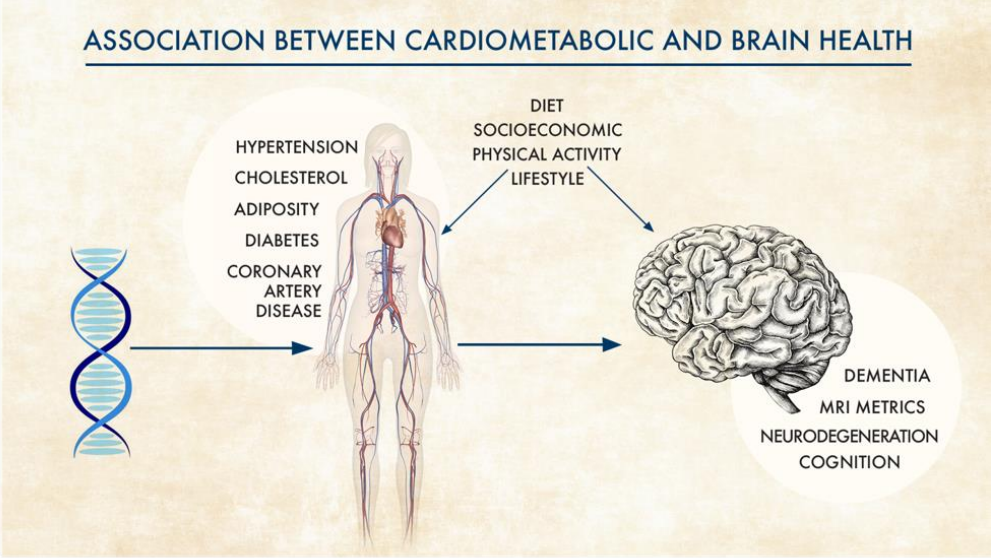

Figure 1.2 Overview of risk factors implicated in literature

1.4 Limitations of current research

To recap, vascular risk factors play a large role in cognitive decline, and anatomical substrates have been well documented in the literature. There is also growing evidence that the use of polygenic risk scoring can improve prediction specifically for CVD and CMD risk and events. Current research on genetic and cardiometabolic contributions to brain health is extensive and has been studied for a number of years. However, there are some limitations within the literature at present which the present thesis aims to address:

- A significant methodological limitation of current research in this area is issues with sample size, particularly as studies with both MRI and genetic data are at risk of being underpowered to detect small changes in those who have not yet developed clinical symptoms. Due to lack of sample size, analyses including blood, genetic and MRI measures have not been extensively studied together in the same cohort.

- A large source of bias and spurious association in the current literature have been due to poor controlling for confounders – variables that explain an

association that has not been considered or accounted for. The majority of study designs studying cognitive or brain MRI metrics in the context of cardiometabolic health have not gathered sufficient information on confounders such as smoking, BMI, medications or other conditions. The UK Biobank resource however has a wide variety of biological, lifestyle, imaging and genetic data that allows for variables to be adjusted for and considered. One objective of this work is to address this gap within the literature by considering external and environmental variables within association analyses between complex phenotypes.

- Currently there is a lack of understanding of how to best stratify and carry out disease risk assessment for individuals who are most at risk of cognitive decline. For this to be incorporated, more research is needed in order to reach a consensus regarding relevant phenotypes and biomarkers. This thesis aims to make a contribution to the evidence.

1.5 Thesis research aims

The present thesis aims to contribute to the current literature by addressing the following research aims:

1. Contribute to understanding of cardiovascular contributions to brain health

Chapter 3 addresses this research aim by comparing brain MRI of individuals varying in number and type of cardiometabolic condition. This was the first study to compare such groups with healthy controls in a cross-sectional design using the same cohort to our knowledge. The aim of this was to examine whether some cardiometabolic comorbidities would reflect worse brain MRI than others. Aim 1 of this thesis was also examined in chapters 4 and 6. Chapter 4 compared differences and associations between genetically elevated lipoprotein A and blood lipoprotein A, a well-established cardiovascular risk factor and predictor, in the context of brain MRI and cognition. Chapter 6 studied an

Alzheimer's genetic risk score, previously shown in chapter 5 to correlate with relevant brain MRI phenotypes, to look at blood biomarker profiles in individuals at risk of poorer brain phenotypes, including biomarkers implicated in cardiometabolic dysfunction.

2. Determine potential role of genetic risk factors on the brain and (concurrent) blood biochemistry in healthy adults

Chapter 4 looked at elevated lipoprotein a (LpA) which is a well-established risk factor for CVD. This chapter aimed to determine whether polygenic risk of elevated LpA may be a valuable genetic instrument for this phenotype and a meaningful risk factor for brain health. In chapter 5, a genetic risk score was created for Alzheimer's dementia which was used to look at associations with brain health, i.e., MRI and cognitive measures, while chapter 6 examined the role of this genetic risk score in physical health, i.e., circulating blood biomarkers.

1.6 Thesis overview and structure

The present thesis includes one methodology chapter (chapter 2) and four analytical chapters (chapters 3-6), which examine the genetic and cardiometabolic contributions to brain health in healthy midlife adults. Chapter 7 provides a discussion and overview of the findings.

The first analytical chapter, 3, was based on evidence that cardiometabolic diseases (CMD) are associated with brain health metrics and may interact with one another, in part due to overlapping mechanisms which accelerate neurodegeneration. However, there is little evidence differentiating the ways in which separate conditions present in the brain, particularly how common comorbidities may differ from one another in terms of brain MRI. Outcomes of CMD multimorbidities have not been well studied, which chapter 3 of this thesis aimed to examine. This was studied using a cross-sectional analysis of 10, 302 UK Biobank participants in midlife with CMDs vs healthy controls. Those with CMDs were split

into 8 mutually exclusive groups based on condition, e.g., hypertension, or type 2 diabetes, and also by number of conditions, e.g., 2 conditions.

Chapter 3 did not find meaningful differences in brain MRI findings between groups, i.e., brain MRI metrics were not more associated with some CMD groups more than others. For this reason, chapter 4 turned to investigating specific markers of cardiometabolic or cardiovascular health which may elucidate changes occurring to the brain in midlife. Chapter 4 investigated the role of a mechanistic blood biomarker: lipoprotein A (LpA). Both blood and genetically elevated LpA are reliable risk factors for CVD, and a substantial risk factor for neurodegeneration and associated conditions such as AD. However, little is known in comparison about the extent to which risk factors for CVD, such as LpA, influence pathology in the brain. Chapter 4 examined whether LpA was associated with brain health measures in 32, 790 individuals by using elevated blood LpA and genetically elevated LpA with brain phenotypes, including white matter structure, brain volumes and cognitive performance. This chapter found both blood and genetically elevated LpA were associated with poorer WM integrity (mean diffusivity).

Chapters 3 and 4 focused on studying the ways in which cardiovascular and cardiometabolic factors are associated with poorer brain health. Chapter 4 found evidence to suggest that genetic risk of elevated LpA, and elevated blood LpA may be variables of interest when examining brain health. Chapter 5 used similar PGR scoring methods to chapter 4 to investigate AD as one of the most common and widely studied phenotypes of neurodegeneration and cognitive decline. In this chapter, we looked at whether the genetic risk of AD, including genes implicated in CVD, can better reveal differences in brain MRI and cognition between healthy individuals in the same cohort. This chapter analysed whether genetic risk (~6 million variants) of AD in 32,790 healthy individuals was associated with brain structure and cognitive abilities. This chapter introduced brain MRI phenotypes such as hippocampal subvolumes, known substrates of dementia and cognitive decline. This was the largest study at the time to study the association of PGR of AD and brain metrics.

Chapter 5 provided evidence that the PGR for AD may indicate the earliest signs of pathology in terms of brain MRI before cognitive deficits were apparent. For this reason, the final analytical chapter, chapter 6, used this PGR of AD to investigate whether genetic risk of AD was associated with 30 blood biomarkers in around 500, 000 individuals. We found differences amongst blood biomarkers when separating males and females, where there was evidence for sex interactions with genetic risk of AD.

Chapter 2: Methodology

2.1 UK Biobank

The UK Biobank is the largest long-term biobank study in the UK, with ~502,000 participants having baseline data collected in the first instance between 2006 and 2010. Baseline data includes demographic, lifestyle, genetic, medical and cognitive data but not imaging. Participants at baseline were between 40-69 with approximately an even number of males (46%) and females (54%). This age range was chosen with a follow up period of 10 – 20 years in order to observe onset of disease from early years (Sudlow et al., 2015). All participants gave full informed consent to the NHS Research Ethics Service at baseline (Ref 11/NW/0382) and study sites for UK Biobank data collection include 22 UK assessment centres with bases in Stockport, Greater Manchester and Scotland.

UK Biobank's first repeat assessment was carried out between 2012 and 2013 in which ~20, 000 participants returned for follow up assessments including medical, lifestyle and cognitive data. A second repeat assessment was carried out beginning in 2014 in which imaging data was introduced for the first time. Lifestyle, imaging and cognitive data was collected in the second repeat assessment and all imaging data used in this thesis is from this second repeat assessment. The UK Biobank is aiming to repeat further follow up assessments at 2-3-year intervals with the aim of scanning 100, 000 subjects by 2022. As of 2019, the first repeat imaging began (these data are not yet available). Updated retrospective timelines can be found at: https://biobank.ctsu.ox.ac.uk/crystal/exinfo.cgi?src=timelines.

This thesis will use baseline and first imaging visit (2014+) data. Clarification of which UK Biobank assessment points used are specified for each variable.

2.1.1 Ethical statement

First written informed consent was obtained from all participants of the UK Biobank study, which received ethical approval from the Northwest Haydock Research Ethics Committee (REC reference for UK Biobank is 16/NW/0274) and permission for approved researchers to disseminate data and samples was granted

(http://www.ukbiobank.ac.uk/ethics/). Ethical permission for usage of the UK Biobank data for this thesis was obtained under UK Biobank Application 17689 (PI Donald Lyall; Genetic, environmental and lifestyle predictors of brain/cognitive-related outcomes).

2.1.2 Data collection

During the baseline visit between 2006 and 2010 consent forms were signed, then touchscreen questionnaires were presented which included questions about socioeconomic markers, ethnicity, medical and mental health history, lifestyle and general health. The touchscreen questionnaire at baseline also included five cognitive tests. Approximately 20, 000 participants attended a repeat of the baseline assessment visit from 2012 – 2013, which included obtaining information on a participant's health and lifestyle, hearing, physical health and cognitive function. In 2014 imaging visits were undertaken which included brain and body MRI scans with the goal of eventually scanning 100, 000 participants. Cognitive assessments were also administered at the first imaging visits, and lifestyle data was taken. Since 2013, periodic web-based assessments have taken place which include web-based cognitive functioning, mental health, pain, food and diet and work (https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100089).

	Visit station	Assessments undertaken
1	Reception	• Welcome & registration • Generating a USB key for Participants
2	Touch screen Section	• Consent • Touch screen questionnaire • Hearing Test • Cognitive function tests (Shape, Pairs, Fluid Intelligence, Snap)
3	Interview & blood pressure	• Interviewer questionnaire • Blood pressure measurement • Measurement of arterial stiffness (Pulse Wave Velocity)
4	Eye measurements	• Visual acuity • Auto-refraction • Intraocular pressure • Retinal image (OCT Scan)
5	Physical measurements	• Height (Standing and Sitting) • Hip & waist measurement • Weight and Bio-impedance (Body Composition) measurement • Hand-grip strength • Ultrasound bone densitometry • Spirometry (Lung function Test)
6	Cardio (Physical fitness)	• Exercise ECG (Cycling)
7	Sample collection & exit	• Blood samples collected • Urine sample sought

Figure 2.1 Chronological sequence of assessments taken during UK Biobank visits

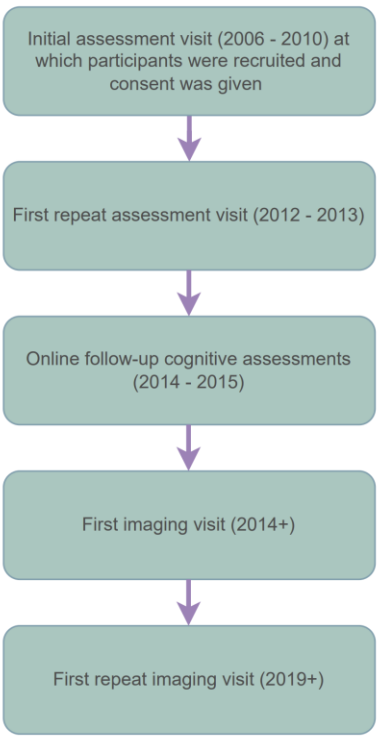

Figure 2.2 Data collection instances

2.2 Sociodemographic and covariate data

Demographic and covariate characteristics used for this thesis include age, ethnicity, sex, education, townsend deprivation score, smoking and BMI. Instance at which data was taken from is specified below.

Age was included in all analyses. For chapters 3, 4 and 5 which used imaging data, age was coded as the age each participant was when imaging data was collected (Data-Field 21003 “age when attended assessment centre”). For chapter 6 where imaging data was not used, age was taken from the baseline measure: Data-Field 21022.

Ethnicity when included as a covariate was taken from baseline as a self-reported variable (Data-Field 21000). However, in analyses that included genetic analyses, only those of white British ancestry were included which was determined by genetic QC carried out by UK Biobank where those who self-reported 'White British' and had similar genetic ancestry based on a principal components analysis of the genotypes were included as white British.

Sex was included as self-reported dichotomous variable from baseline in which analyses with genetic QC excluded those who had a self-report and genetic sex mismatch.

Education (Data-Field 6138) was based on self-reported highest qualification received and then was dichotomised into “University degree” or “less”. A-Level or equivalent, GCSE or equivalent, CSE, NVQ, HND or equivalent, other professional qualification or none of the above were all dichotomised into “less than university degree”. Self-reported education was included in all analyses. Chapters 3, 4 and 5 used data from the first imaging visit. Chapter 6 used baseline data.

Townsend deprivation index (Data-Field 189) (Townsend, 1987) was derived from postcode of residence for which a socioeconomic deprivation score was calculated based on national census. The Townsend index is a measure of deprivation in a

population which considers four variables: unemployment (as a percentage of those aged over 16), non-car ownership (as a percentage of all households), non-home ownership (as a percentage of all households) and household overcrowding. It is possible for a Townsend score to be calculated for any area in which there is information on the four variables. The UK Biobank Townsend scores combined data from census output areas, this was based on around 125 households per area in England and Wales and around 50 households per area in Scotland. Townsend scores were used from baseline, however, there was an option to amended postcode of residence by the participant upon arrival at each assessment.

Smoking was used as a covariate in all analyses as either pack/years or ever vs never. Pack years was an average number of packs smoked per year proportional to lifetime exposure as measured by age. Pack/years (Data-Field 20162) was calculated as: Number of cigarettes per day / 20 * (Age stopped smoking - Age start smoking) the output was then taken and calculated as such: Pack years / (Age at recruitment - 16) to provide an approximate lifetime exposure. This measure was available for all instances, in which the most recent or relevant calculation was included for all analyses, similarly to age.

Body mass index (BMI) (Data-Field 21001) was taken at baseline and was measured by trained research staff in which participants removed shoes and heavy outer clothing. Weight was measured, to the nearest 0.1 kg, using a Tanita BC-418MA body composition analyser and height using a Seca 202 height measure. BMI was derived from: weight (kg)/(height (m) height (m)) and used in analyses as a continuous measure in Kg/m2. BMI was taken at each visit, where chapters 3, 4 and 5 used BMI data from the first imaging visit. Chapter 6 used baseline BMI.

2.3 Cognitive data

The touchscreen questionnaire included a battery of cognitive tests that were designed specifically for UK Biobank. Time to complete the main five cognitive assessments was approximately 15 minutes (numeric memory was added later and

has subsequently been removed from the battery). Cognitive measures were taken at baseline, and sub-samples of participants were assessed in subsequent follow ups which included the repeat, first imaging and repeat imaging. Several of the cognitive function tests administered via touchscreen during the initial assessment visit were administered at subsequent instances as web-based questionnaires and participants were invited to complete them remotely. Three tests (fluid intelligence, numeric memory and prospective memory) were introduced in the final two years of baseline recruitment, with numeric memory subsequently being removed for time. Therefore, instances at which cognitive data was taken from vary and are clarified below after descriptions of each cognitive measure in figure 2.7.

2.3.1 Fluid intelligence

Fluid intelligence (Data-Field 20016) consisted of 13 logic questions and was defined as “the capacity to solve problems that require logic and reasoning ability, independent of acquired knowledge” in which participants had 2 minutes to complete as many of the 13 logic questions. Cronbach alpha coefficient for these items has been reported at 0.62 (Lyall et al., 2016). Logic questions included a numeric addition test, identification of the largest number, word interpolation, positional arithmetic, familial relationship calculation e.g. "If Truda's mother's brother is Tim's sister's father, what relation is Truda to Tim?", conditional arithmetic, synonym, chained arithmetic, concept interpolation, arithmetic sequence recognition, antonym, square sequence recognition and subset inclusion logic. Fluid intelligence score was calculated as sum of correct answers.

2.3.2 Prospective memory

For prospective memory (Data-Field 20018) participants were asked to engage in a specific behaviour later in the assessment. "At the end of the games we will show you four coloured shapes and ask you to touch the Blue Square. However, to test your memory, we want you to actually touch the Orange Circle instead".

Participants were given four attempts, and this was coded in the data as either 0 or 1 depending on whether the participant gave a correct answer or not.

2.3.3 Trail making (A+B)

The trail making task (TMT; Data-Fields 6348 and 6350) was introduced in the first imaging visits. For TMT part A, participants were instructed to connect numbers from 1 – 25 in ascending order as quickly as possible. Digits were quasi-randomly distributed on the screen. For TMT B, participants were required to select both letters (A-L) and numbers (1-13) in alternating and ascending order e.g. 1 A 2 B 3 C etc. The intervals between two touching points were taken in seconds and summed for both A + B for the TMT score in analyses for this thesis.

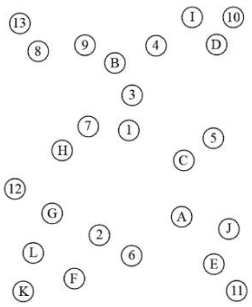

Figure 2.3 Example screen shown during trail making task

2.3.4 Symbol digit substitution task

Number of symbol digit matches made correctly (Data-Field 20159) was collected from the digit symbol substitution task in which individuals were required to match symbols to numbers using a key at the top of the screen. Participants had a brief practice consisting of 8 substitutions before the live timed segment began in which the number of correct matches made within 2 minutes.

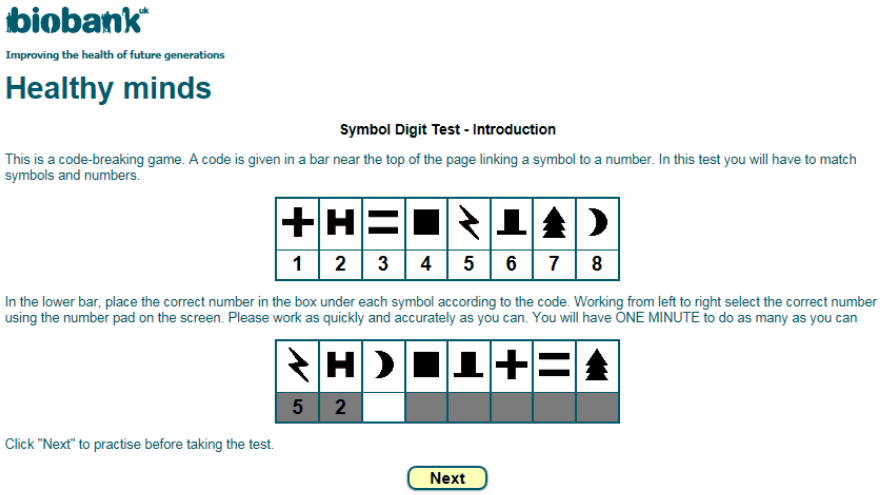

Figure 2.4 Example screen shown during symbol digit substitution task

2.3.5 Reaction time

The Reaction time (RT) (Data-Field 20023) task consisted of 12 rounds of the card game 'snap' in which participants were shown two cards at the same time on the screen and instructed to press the button on a button box as quickly as possible when the symbols on the cards matched, in a 'go-no-go' manner. The RT score for each participant was calculated as mean response times for correctly identified matching pairs in milliseconds and Cronbach's alpha for this task reported at 0.85 (Hagenaars et al., 2016). RT was positively skewed; therefore this variable was log transformed and standardised for all analyses.

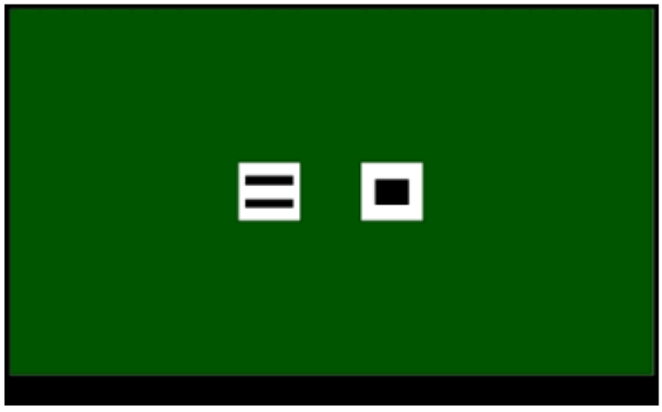

Figure 2.5 Example screen shown during reaction time task

2.3.6 Numeric memory

Numeric memory (Data-Field 4282) was tested by showing participants a string of numbers and asking them to recall from memory via a numeric keypad, the number increased by one and numeric memory score was recorded as longest string length correctly recalled. Each string was shown on screen for 2000 milliseconds + 500 milliseconds multiplied by the string length. The test began with two digits and advanced to a maximum number of 12 digits. The test was discontinued for two incorrect responses at a string length of three or more, or five incorrect responses at a string length of two.

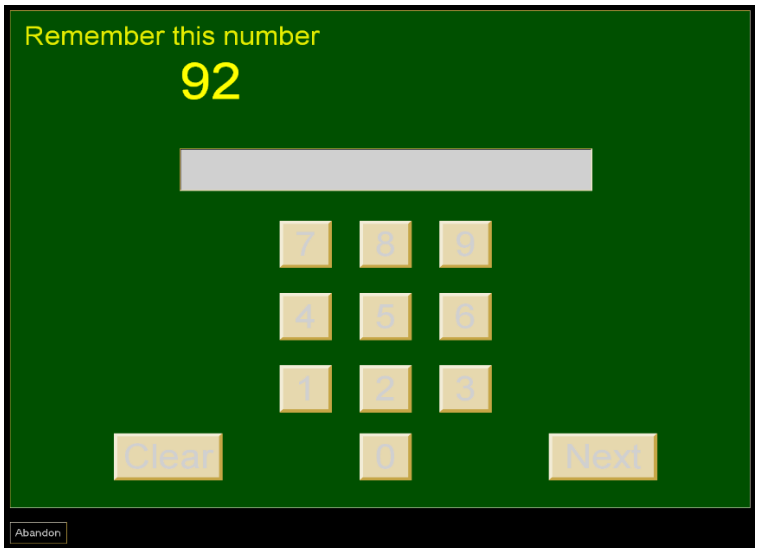

Figure 2.6 Example screen shown during numeric memory task

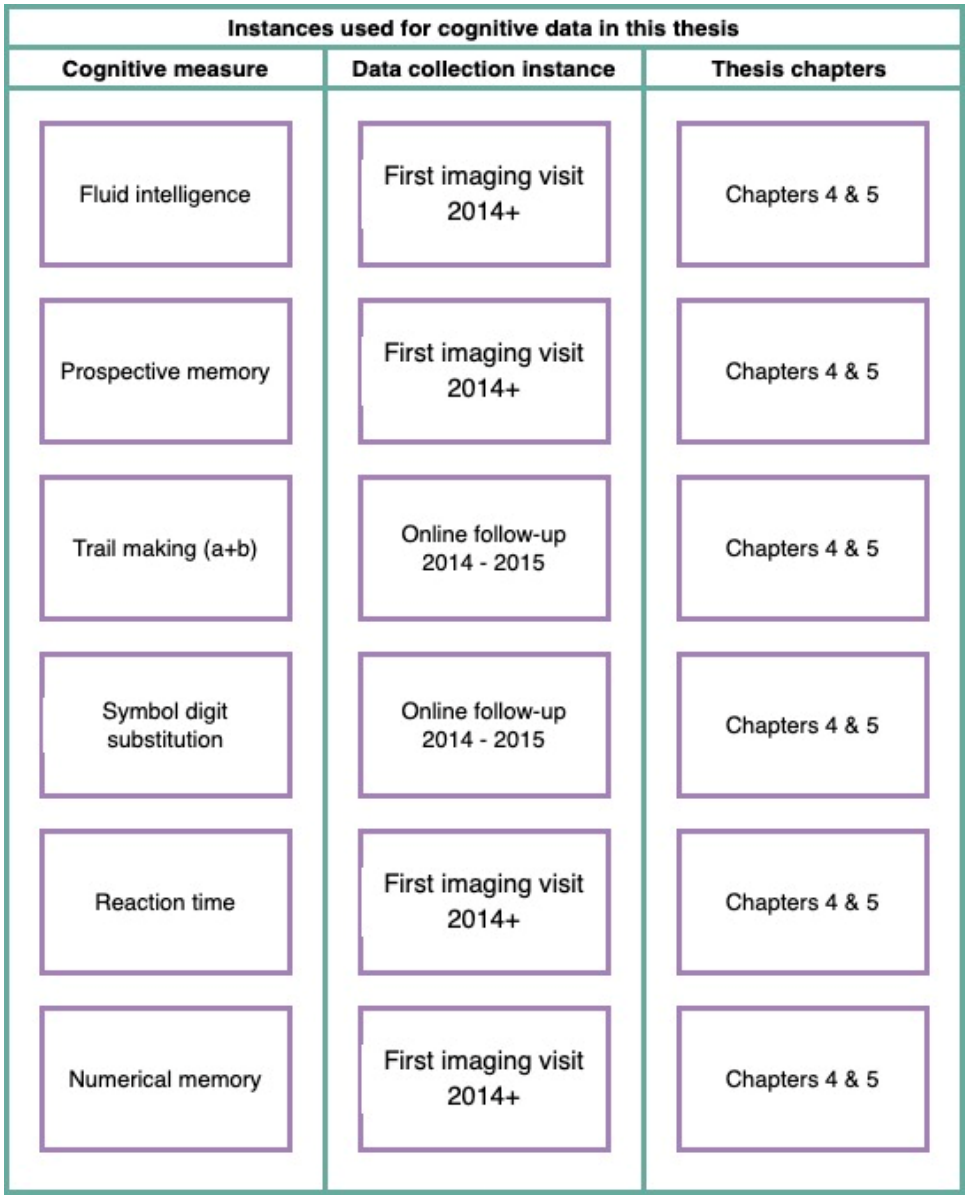

Figure 2.7 Cognitive measure collection instances and corresponding chapters

2.3.7 Reliability and validity of UK Biobank cognitive assessments

The psychometric properties of UK Biobank cognitive tests have been studied by Fawns-Ritchie and Deary (2020), Hagenaars et al. (2016) and Lyall et al. (2016). Fawns-Ritchie and Deary (2020) carried out analyses with 160 participants recruited at The University of Edinburgh (age range 40 to 80) who took the UK Biobank cognitive test battery and corresponding reference tests. The reference tests used were well-validated, standard cognitive tests intended to measure the same

underlying cognitive domains as those used by the UK Biobank. A subset of participants (n= 50) returned to retest on the UK Biobank assessments 3 to 6 weeks following the initial assessment. This was conducted to study the short-term test-retest reliability - this was of interest as the UK Biobank's cognitive tests are intended to assess longitudinal change. Fawns-Ritchie and Deary aimed to investigate whether UK Biobank tests and their equivalent reference tests would correlate higher than tests assessing different domains, as well as the short-term stability of the UK Biobank tests. They also investigated whether a g component was present in the UK Biobank and corresponding reference tests, and whether UK Biobank and reference tests correlated with one another. A g component is a psychometric construct that calculates a composite score from the inter-correlational structure of cognitive tests. This is calculated through a principal component analysis, where the first unrotated principal component typically accounts for about 40% of variance in a wide range of different cognitive tests (Deary, 2013; Deary et al., 2019). The authors identified a g component in the original UK Biobank tests and the reference tests. They also found that, although the UK Biobank tests were brief, they did correlate with reference tests purported to test the same cognitive ability. This supports the UK Biobank cognitive tests as having good concurrent validity.

Symbol digit substitution is considered a processing speed task, whereas TMT is a test of visual attention and task switching. Fawns-Ritchie & Deary (2020) found that TMT and symbol digit substitution both showed good concurrent validity with respective reference tests ($r = 0.50$, $r = 0.64$, respectively), and with tests that required speed. Hagenaars et al., (2018) also examined TMT in the UK Biobank. They investigated whether cognitive abilities required for TMT performance were distinct from other cognitive domains. This was of interest due to TMT requiring the use of several cognitive processes including attention, visual searching, psychomotor speed, abstraction, flexibility and working memory. They found that TMT was both phenotypically and genetically associated with general cognitive function and processing speed. This would support that processing speed and working memory account for age effects on TMT.

Reaction time (r = 0.52) and numerical memory (r= 0.43) also showed good concurrent validity (Ritchie & Deary, 2020). Lyall et al., (2016) also found longitudinal stability in reaction time (intraclass correlation = 0.57; 95% CI = 0.56 - 0.58). However, it is not possible to further comment on reliability or validity of the UK Biobank numerical memory assessment as this was not included for participants at baseline.

Fawns-Ritchie & Deary (2020) examined the UK Biobank prospective memory test, which correlated moderately with other memory tests, as well as tests of executive function (r = 0.30) and reasoning (r = 0.41). However, it did not correlate highly with the chosen reference test (r = 0.22). It is important to note that there may have been a ceiling effect for this measure as it was recorded as either completed or not completed, which may contribute to low reliability and validity.

Lastly, although Fawns-Ritchie & Deary (2020) did not include a reference test for fluid intelligence to assess concurrent validity, they found that it was negatively correlated with age and correlated with other reference tests of working memory and non-verbal reasoning. This suggests that UK Biobank's measure of fluid intelligence shows some validity. Lyall et al.'s (2016) study investigating longitudinal stability over an average period of four years found two-way intraclass correlations were high for fluid intelligence (ICC = 0.65, 95% CI 0.63 - 0.67).

2.4 Blood sample collections

45 ml (6 tubes) of blood was collected from each participant at baseline. A total of 30 blood biomarkers were chosen due to their relevance with diseases and risk factors (https://www.ukbiobank.ac.uk/media/oiudpjqa/bcm023_ukb_biomarker_panel_website_v1-0-aug-2015-edit-2018.pdf). More details about biomarker collection and processing are provided in the relevant chapters.

2.5 Genotyping and imputation

Participants provided a blood sample at the baseline assessment for which two similar genotyping arrays were used. Genotyping, imputation and QC was carried out centrally by UK Biobank for which documentation can be found here: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/genotyping_qc.pdf and https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/impute_ukb_v1.pdf. QC data are available for download and include data on Hardy-Weinberg equilibrium, pairwise relatedness of individuals, mismatch between reported and genetic sex, and probable white British ancestry.

The majority of the UK Biobank participants (~450,000) were genotyped using the Applied Biosystems UK Biobank Axiom Array which directly measured 825, 927 markers. The remaining ~50, 000 participants were genotyped using UK BiLEVE Axiom array by Affymetrix which included 807, 411 markers. These arrays have over 95% content in common and they included SNPs chosen because of potential associations with diseases and health-related phenotypes. They were also chosen because of their common (>5%) and low (1 - 5%) minor allele frequencies in European populations. Imputation was based upon a merged reference panel of 87,696,888 biallelic variants on 12,570 haplotypes constituted from the 1000 Genomes Phase 3 and UK10K haplotype panels. Principal components analysis was conducted by UK Biobank to identify ancestral population structure within the cohort (Price et al., 2006), and 8 principal components were used as a covariate for analyses to avoid spurious association due to individuals sharing genetic ancestry. This thesis elected to use 8 principal components because there is evidence that the majority of variance is explained by 5 where 8 is a common approach (Galinsky et al., 2016).

2.5.1 Polygenic risk scores

Two polygenic risk scores were created using LDpred for the work of this thesis. External genome-wide association study (GWAS) summary statistics and UK Biobank genotype data were used and discussed in relevant chapters. LDpred (Vilhjálmsson

et al., 2015) is a Bayesian approach to polygenic risk scoring those accounts for linkage disequilibrium through effect sizes. We used the infinitesimal model of LDpred to create polygenic risk (PGR) scores as other polygenic risk scoring methods using many small effect SNPs may lead to a more inaccurate PGR score and thus increase type 1 or type 2 error rates in estimating overall genetic effect size. We therefore elected for a method that uses more accurately estimated SNP effect sizes. LDpred reweights raw effect sizes by their linkage disequilibrium using a reference panel of 1000 unrelated people, for this thesis we used individuals from the UK Biobank cohort who passed genetic QC but did not have imaging or biomarker data and therefore were not included in analyses.

2.6 Imaging metrics

Imaging data in UK Biobank was collected for the first time in 2014 – 2019 for brain (3T MRI), as well as heart (DXA), abdomen (DXA), bone (DXA) and carotid arteries (ultrasound). UK Biobank brain MRI data includes raw T1, T2, FLAIR, diffusion MRI (dMRI) and both resting state fMRI (rfMRI) and task fMRI (tfMRI) (protocol: http://www.fmrib.ox.ac.uk/ukbiobank/protocol/V4_23092014.pdf and documentation: http://biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf). All brain MRI data was acquired on a Siemens Skyra 3T with a 32-channel head coil in which T1, T2 (111mm resolution) and FLAIR (1.0511mm resolution) images were acquired in sagittal orientation. dMRI acquisition comprised a spin-echo echo planar sequence (Miller et al., 2016). Image derived phenotypes (IDPs) were processed through a pipeline (Alfaro-Almagro et al., 2018) which can be found on github (https://git.fmrib.ox.ac.uk/falmagro/UK_biobank_pipeline_v_1), and were made available as 2,501 separate IDPs. There are a total of 25 volumetric IDPs derived from T1 scans including subcortical structures and whole brain volumes for white and grey matter. T2 and FLAIR images provided anatomical phenotypes and white matter hyperintensity measures. dMRI IDPs provided information about tissue microstructure including directionality of water molecules within brain tissues, giving information about white matter tract integrity. Processed dMRI phenotypes include 432 data-fields (category 134) including intracellular volume fraction (ISOVF), intracellular volume fraction (ICVF), orientation dispersion (OD), fractional anisotropy (FA) and mean diffusivity (MD).

All MRI volumetric measures were in mm3 and include whole brain volume (Data-Field 25009), volume of grey matter (Data-Field 25005), volume of white matter (Data-Field 25007), volume of white matter hyperintensities (Data-Field 25781), whole left (Data-Field 25019) and right (Data-Field 25020) hippocampal volumes. Hippocampal subfields were also used, including: right hippocampal body (Data-Field 26661), head (Data-Field 26662) and tail (Data-Field 26642), and left hippocampal body (Data-Field 26639), head (Data-Field 26640) and tail (Data-Field 26641).

IDPs used in this thesis were normalised for head size either manually or as part of UK Biobank pre-processing, however, a study by Lyall et al., (2013) compared methods to adjust for brain size when looking at hippocampal volumes and found that differing normalisation methods may influence associations found with brain volumes.

2.7 Inclusion criteria

Participants were excluded from all analyses if they self-reported (category 2406) the following conditions from the UK Biobank touchscreen questionnaire: Brain cancer, brain or subarachnoid haemorrhage, cerebral aneurysm, cerebral palsy, chronic/degenerative neurological problem, dementia/Alzheimer's disease, encephalitis, epilepsy, meningitis, motor neurone disease, multiple sclerosis, neurological injury/trauma, Parkinson's disease, spina bifida, stroke, transient ischaemic attack.

2.8 Cohort demographics and characteristics

When looking at the UK Biobank cohort as a whole, there were approximately 502,690 individuals aged between 37 to 73 years old with the vast majority (500,205) aged between 40 and 69. 54% of participants at baseline were female, and 88% identified as white British. Fry et al., (2017) studied the demographics of the UK Biobank cohort in comparison to the 2011 UK census report and reported that the UK Biobank is not a representative sampling of the UK due to several differences. The study found evidence of a healthy volunteer bias in which UK Biobank participants are more likely to smoke less, identify as white, own their property, have lower BMI and have fewer self-reported health conditions (cardiovascular disease, stroke, chronic kidney disease, cancer and respiratory disease) in comparison to the general population. Other studies have found that the UK Biobank may not adequately represent the most deprived areas within the UK (Pham et al., 2018; Batty et al., 2019; Wilkinson et al., 2019; Lyall et al., 2021) which is problematic as individuals in more deprived areas arguably may be most affected by health conditions being studied via the UK Biobank. This finding is

further exacerbated in individuals who underwent imaging assessments as there is evidence that they are healthier in comparison to the greater UK Biobank cohort. This effect was seen for the imaging subsample in a range of cardiometabolic, cognitive and mental health phenotypes. This may in part be due to exclusions such as MRI contradictions such as stents or pacemakers (Lyall et al., 2021). Demographics of UK Biobank data may have also changed over the years due to participants withdrawals, in which the longitudinal studies using follow up data may be even less generalisable to minority and socioeconomically deprived individuals.

Chapter 3: Associations between cardiometabolic multimorbidities and brain MRI metrics in UK Biobank

3.1 Introduction

Cardiometabolic conditions, collectively, are the leading cause of preventable death worldwide. There is a large body of literature supporting the role of cardiometabolic disease in the decline of both physical and brain health due to detrimental effects on the heart or blood vessels (Rao 2018; Kivimäki et al., 2019; Khan et al., 2016). There is also an increasing awareness of the overlap between cardiometabolic disease (CMD) and brain pathologies, in which CMDs have been consistently associated with decreased cerebrovascular health and increased damage to white matter in the brain (Gupta et al., 2018; Wardlaw et al., 2015; Wang et al., 2015). Subsequently, there has been an increase in studies identifying biological mechanisms that contribute to the development of brain pathologies as a result of cardiometabolic conditions.

Over time CMDs gradually change vascular structure and function, and whilst some mechanisms overlap, there are hallmark pathological changes of individual CMDs. Hypertension (HT) is a CMD characterised by persistently elevated blood pressure, which increases shear stress burden on the vascular system, causing morphological changes to blood vessel walls (Thomas et al., 2016). In contrast, mechanisms and risk factors underlying type 2 diabetes (T2D) are understood to contribute towards oxidative stress and endothelial dysfunction (Stehouwer et al., 2015). coronary artery disease (CAD) on the other hand is closely linked to atherosclerosis and is characterised by cholesterol plaque deposits causing vessel blockages. Atherosclerosis is typically a consequence of high triglyceride levels, high blood pressure, high glucose, insulin levels or hypercholesterolemia (Zachariah et al., 2018), leading to CAD, which is associated with plaque build-up and the consequent narrowing of vessels which decreases blood flow (Zachariah et al., 2018).

Although there are characteristic disease markers for each cardiometabolic condition described above, some mechanisms involved in each disease trajectory can overlap between conditions. These include the dysregulation of homeostasis in lipid metabolism and insulin, which can contribute to overall brain health via direct

cerebrovascular damage and inflammation – effects that may accelerate brain injury. Comorbidity of T2D, CAD and HT, in particular, may in part be explained by common metabolic activity. For example, chronic hyperglycaemia in T2D may lead to additional vascular complications or conditions. This can occur through increased production of advanced glycation end products (AGEs). AGEs have a range of pathological effects including vascular permeability, oxidising low density lipoproteins, and endothelial dysfunction resulting from insulin resistance which is a known precursor to atherosclerosis due to metabolic disruption of vasoconstriction. Another mechanism of T2D which is thought to initiate the development of HT is through the abnormal metabolic state associated with diabetes which decreases the blood vessel's ability to vasodilate due to arterial stiffening, initiating atherosclerosis and consequent high blood pressure. Diabetes is also known to increase the amount of fluid in the body, contributing to high pressure in the vessels. Additionally, hyperglycaemia and insulin resistance in diabetes change how the body manages insulin which influences blood pressure as insulin is a pleiotropic hormone, playing an important role in the development of HT.

Although HT and T2D are an example of how CMDs can share common mechanisms, it is not yet understood why some individuals develop multiple diseases and whether some types of multimorbidity are more detrimental than others. The comorbidities between T2D, CAD and HT may be due to pathology of one cardiometabolic condition initiating pathology for another, which can occur in several ways.

3.1.1 CMD and brain MRI phenotypes

Such biological mechanisms associated with cardiometabolic conditions have been heavily implicated in cerebrovascular disease and brain health with tissue-specific effects (Debette et al., 2019). The role of cerebrovascular health in neurodegeneration is supported by many studies investigating causes of white matter degeneration (Biesbroek et al., 2017; Schilling et al., 2014; Gupta et al., 2018; Wardlaw et al., 2015). These studies have provided evidence of a CMD and

brain health link where cardiometabolic mechanisms are thought to play a key role in the aetiology of white matter neurodegeneration and a higher burden of white matter hyperintensities (WMH). A longitudinal cohort study by the Mayo Clinic Study of Aging found that baseline hypertension, hypertension in midlife, and fasting glucose were predictive of WMH progression over time (Scharf et al., 2019). Other studies have similarly found that the prevalence of WMH increases with vascular risk factors, including hypertension (Dufouil et al., 2001; Van Dijk et al., 2004) and chronic hyperglycaemia (Ferguson et al., 2003). Wang et al. (2015) found that vascular risk factors such as high blood pressure were associated with reduced brain white matter integrity due to increased stress to microvessels, namely cerebral small vessels, and were able to predict consequent brain pathology and cognitive decline.

HT and T2D are one example of common cardiometabolic comorbidity associated with poorer vascular and brain health metrics (Petrie et al., 2018; Yahagi et al., 2017; Schmidt et al., 2017; King et al., 2014; Lastra et al., 2014). One study found that CAD increased the risk of HT and dyslipidaemia (Murray et al., 2018). The San Antonio heart study found that 85% of type 2 diabetics were also hyperintense. Around half of the hypertension patients showed evidence of insulin resistance, i.e., impaired glucose tolerance (Lorenzo et al., 2003), suggesting an overlap and possible interaction in metabolic pathways. These conditions may share common metabolic pathways contributing to the development of disease states, including adiposity, inflammation from adaptive immune response, upregulation of the angiotensin-aldosterone system, abnormal sodium handling and insulin resistance (Petrie et al., 2018). A study with The Action to Control Cardiovascular Risk in Diabetes (ACCORD) cohort showed that treating comorbid conditions in diabetes patients, such as reducing high blood pressure, was more effective in reducing vascular complications of diabetes compared to targeting blood glucose levels (Mannucci et al., 2013). This suggests that comorbidity is an important factor to consider when seeking to reduce cardiovascular risk and associated outcomes, including compromised brain health. These findings also suggest type of comorbidity is important to consider within multimorbidity.

One randomised controlled trial of individuals (n=9, 361) over 50 years of age who had hypertension but no history of T2D or stroke found that when participants received intensive intervention treatment for HT, they were less likely to decline cognitively at follow up (median=3.34 years) when compared to those who received a standard treatment for blood pressure (Mayor et al., 2019). The study authors concluded that intensive control of blood pressure (goal of <120 mm Hg systolic blood pressure) significantly reduced the risk of cognitive impairment in which the risk of probable dementia was also reduced (confidence interval: 0.74 to 0.97; p=0.01). This finding not only supports the role of vascular contributions to cognitive impairment but suggests that control of HT may mediate neurovascular damage. The results of this study are consistent with other research, some of which identify age-specific effects of vascular cognitive impairment. King et al. (2014) found that the relationship between increasing cardiometabolic comorbidity and WMH increased significantly after 50 years of age, suggesting detrimental effects of cardiometabolic comorbidity may not present until later in life. This study, however, only considered WMH volumes for individuals with the specific comorbidity of HT and T2D.

Although there is reliable evidence that mechanisms leading to multimorbidity interact in complex ways, is important to remember that comorbidity can also in part be explained by aetiology resulting from overlapping environmental exposures; examples include nutrition, genetics, air pollution, BMI, effects of genetic risk, as well as the interplay between each disease pathology (Yusuf et al., 2020). While T2D, CAD and HT are associated with white matter damage to the brain individually, the influence of multimorbid T2D, CAD and HT on neurovascular health and cognitive consequences is unclear. Understanding is fragmented and may not consider the combined impact of the conditions. Consequently, the influence of comorbidity on neurocognitive outcomes such as brain MRI is unknown. With a significant incidence of HT, CAD and T2D co-occurring in the population, and individuals living longer with comorbidities, it is crucial to understand potential effects on the brain and contributions to cognitive decline.

3.1.2 Previous UK Biobank studies

The largest single cardiometabolic and brain imaging study (N=9, 722) found that a larger number of vascular risk factors were associated with lower grey matter (GM) and white matter brain volumes and poorer white matter health (Cox et al., 2019). Having a larger number of vascular risk factors was associated with worse brain MRI measures for all metrics (ß range 0.042 to 0.110). WMH were associated with presence of HT (standardised ß = 0.097, p <0.001) and T2D (standardised ß = 0.065, p <0.001) and significantly with increased global atrophy. This study concluded that increased vascular risk factors such as high BMI, high blood pressure and smoking were associated with poorer brain health. However, it did not investigate multiple comorbid conditions and differences between one another. Whilst HT and T2D were considered vascular risk factors in analyses, the coexistence of both conditions and others were not studied, which is the aim of the present research.

One recent UK Biobank study investigated the influence of cardiometabolic comorbidity in relation to cognitive outcome. Lyall et al., 2017 (N= 478, 557) found that cardiometabolic diseases were associated with worse cognitive abilities, and this association was greater with more than one cardiometabolic disease. Other research investigating comorbidity of HT and T2D have reported preliminary findings to support further research into the influence of cardiometabolic comorbidity. It is unclear how comorbidity may affect the developing brain pathologies. This research investigates the effects of HT, CAD and T2D comorbidity on brain MRI measures.

3.2 Methods

3.2.1 UK Biobank

In 2014, around 40, 000 UK Biobank baseline participants were invited and recruited to return for the first brain and body imaging and further follow up assessments. Participants used for this study were from the first MRI imaging dataset of which there were 21, 225 individuals with complete MRI data at time of analyses.

Individuals were excluded for meeting exclusion criteria described below or having incomplete cardiometabolic data, leaving a final 10, 302 individuals that were included in all analyses. Further details about data derived from the UK Biobank, including processing of MRI data, are found in Chapter 2.

3.2.2 Cardiometabolic disease subgroupings

Cardiometabolic diseases were self-reported as part of the UK Biobank health and medical history touchscreen questionnaire. The question (data-field 6150 seen in Figure 3.1) asked participants if they had “Vascular/heart problems diagnosed by doctor” with the option of selecting one or more of the following: “Heart attack”, “Angina”, “Stroke”, “High blood pressure”, “None of the above” and “prefer not to answer”. Participants who selected “Prefer not to answer” or did not respond were excluded from all analyses. Participants who selected “Heart attack” and/or “Angina” received a CAD assignment. Participants who selected “High blood pressure” received a HT assignment. For categorisation of T2D, another question (data-field 2443 seen in Figure 3.2) asked “Has a doctor ever told you that you have diabetes?” with options of “Yes”, “No”, “Do not know” and “Prefer not to answer” in which participants who selected “Prefer not to answer” or “Do not know” were excluded from analyses.

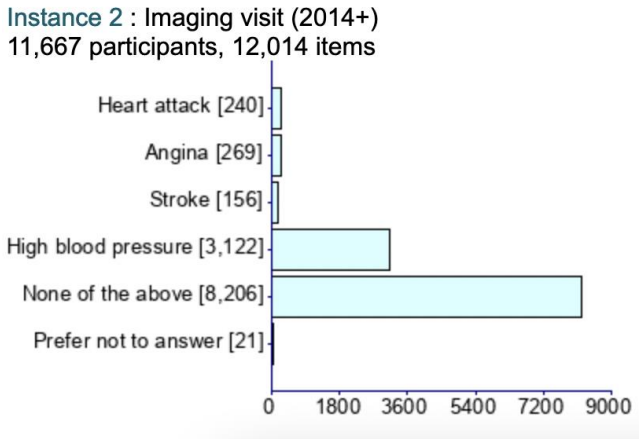

Figure 3.1 Self-reported vascular problems

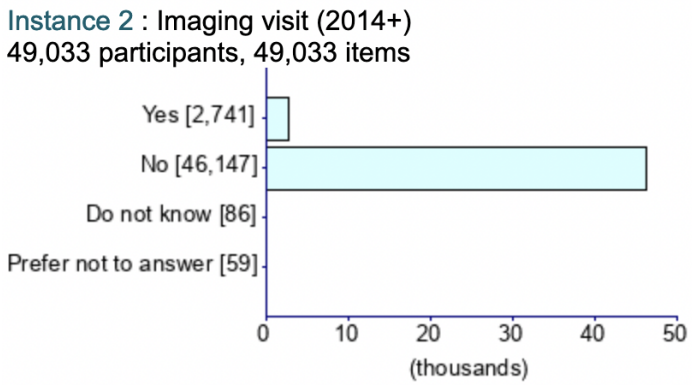

Figure 3.2 Self-reported Type 2 Diabetes

As shown in Figure 3.3, CMD assignments were used to place each participant into one of 4 condition **groupings** or further divided in 8 mutually exclusive **subgroupings**. Subgroups 2, 3 and 4 on the far right included individuals who reported having one cardiometabolic condition. Subgroups 5, 6 and 7 included individuals who reported having exactly two cardiometabolic conditions. Subgroup 8 included individuals that reported having all three conditions i.e., HT, T2D and CAD. All individuals who reported having "None of the above" were placed in subgroup 1 (no cardiometabolic conditions). For analyses, cardiometabolic disease groupings were sometimes collapsed according to the number of cardiometabolic conditions: subgroups 2, 3 and 4 were placed in "one cardiometabolic condition". Subgroups 5, 6 and 7 were placed in "two cardiometabolic conditions". Subgroup 8 was "three cardiometabolic conditions", and individuals reporting no conditions were "No cardiometabolic conditions".

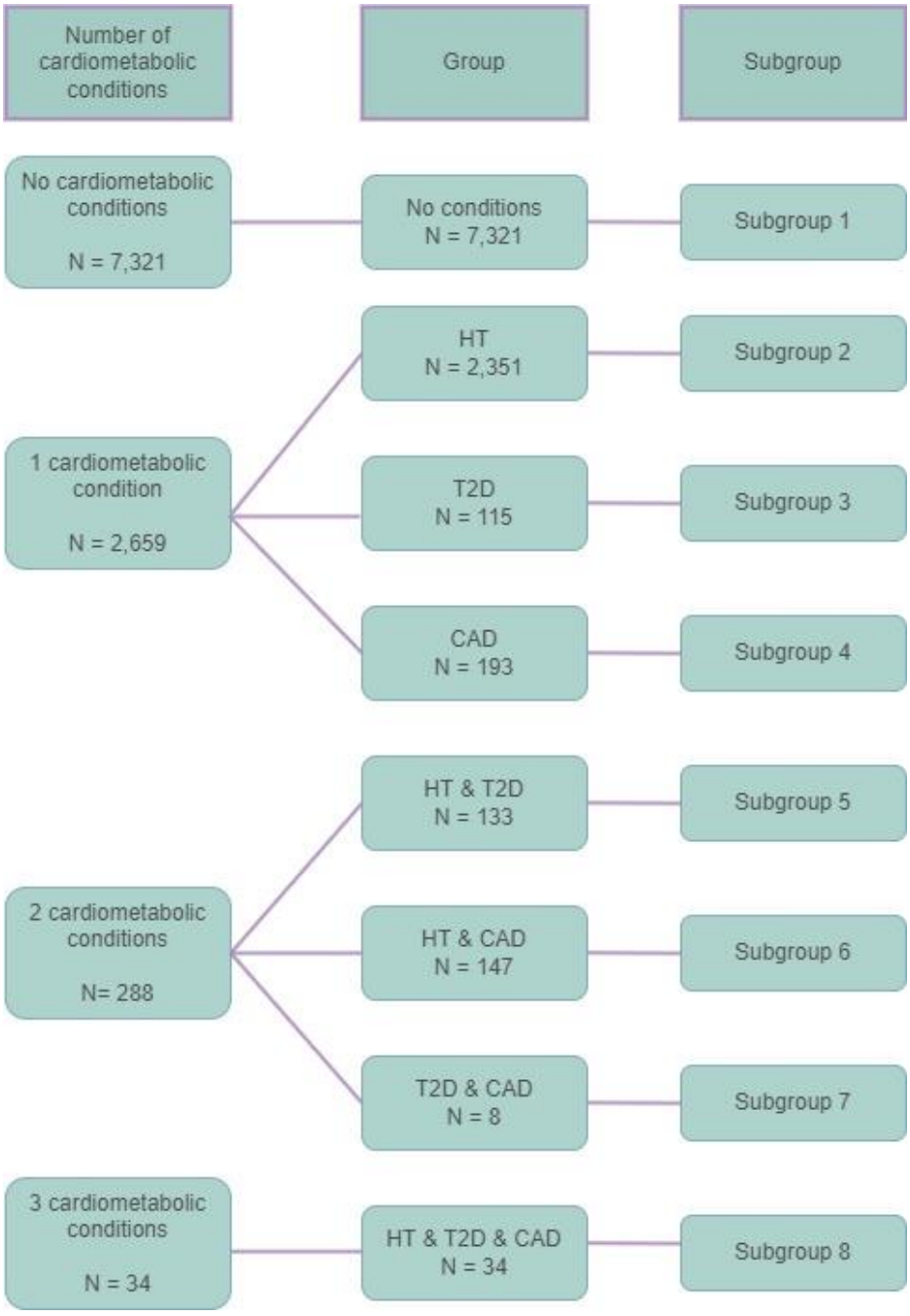

Figure 3.3 Mutually exclusive CMD groupings

3.2.3 Brain MRI Measures

All brain MRI data was acquired on the same 3T Siemens Skyra scanner (Miller et al., 2016). Brain imaging data was processed by UK Biobank and available to researchers as Image Derived Phenotypes (IDPs) (Alfaro-Almagro et al., 2018). IDPs used in these analyses include total brain volume, total grey matter volume and white matter hyperintensity volume which were all measured as mm^3 and normalised for head size. All MRI measures were converted to Z-scores for interpretation and comparison.

3.2.4 Covariates

Covariates included common variables adjusted for within population studies of cardiovascular health (Lyall et al., 2017). Townsend deprivation scores were derived from self-reported postcode of residence and used as a continuous variable in analyses with lower scores indicating more social deprivation. Education was based on self-reported highest qualification received and dichotomised into university degree or less i.e., an individual with an undergraduate degree and postgraduate degree were placed in the same category. Smoking was coded as number of packs smoked per year as a proportion to lifetime exposure and was taken from data collected at the first imaging visit (2014+).

3.2.5 Inclusion criteria

Participants were excluded from analyses if they have reported head injury, brain injury received a diagnosis of dementia, stroke, Parkinson's disease (<5%).

3.2.6 Analyses

When running regressions for number of cardiometabolic diseases, they were grouped and entered as numerical variables (0, 1, 2 or 3) and run in both partially and fully adjusted models. Partially adjusted models were adjusted for age, sex, BMI, and ethnicity. Fully adjusted are adjusted additionally for Townsend deprivation scores, smoking (pack years) and education.

3.3 Results

3.3.1 Demographic statistics

Table 3.1 shows overall descriptive statistics for cardiometabolic groups. The group with no cardiometabolic conditions (subgroup 1) present with lowest age, highest percentage of female, least pack/years smoking and lowest BMI. Subgroup 1 also present with highest total brain and grey matter volumes and lowest overall white matter hyperintensity. There are fewer observations for individuals with three cardiometabolic conditions than other groups (N=34) which is likely underpowered and responsible for skewed descriptive characteristics e.g., 26% female. Percentage female, age, pack/years of smoking, Townsend score, education and BMI all increase in detrimental direction as number of comorbid conditions increases.

The next table, Table 3.2 shows descriptive characteristics of each subgroup with two cardiometabolic conditions where subgroup 7 (T2D & CAD) included 8 individuals, which is likely to be statistically underpowered.

Table 3.3 shows descriptive characteristics of each single cardiometabolic condition show that group 0 show the healthiest metrics. Subgroup 3 (T2D) present with a similar volume of WMH to subgroup 1.

Table 3.1 Cardiometabolic group descriptive statistics by number of conditions

	Cardiometabolic comorbidity groups			
Variable	**No conditions**	**1 Condition**	**2 Conditions**	**3 Conditions**
N	7, 321	2, 659	288	34
Age (M, SD)	60 (7.5)	63.63 (6.3)	65.42 (6.3)	66.21 (5.7)
Sex, N (%F)	4, 155 (56.75)	1,160 (43.66)	89 (30.9)	9(26.47)
Pack Years (M, SD)	17.54 (13.54)	21.31 (16.31)	28.96 (18.58)	28.36 (30.5)
Townsend score (M, SD)	-1.97 (2.53)	-1.93 (2.64)	-1.56 (3.14)	-1.19 (3.11)
Education (% degree)	52%	56%	56%	50%
BMI (Kg/m^2) (M, SD)	26.05 (4.17)	28 (4.64)	30.17 (5.0)	31.88 (5.67)
Total Brain Volume (M, SD)	1514.13 (72.25)	1489.07 (70.03)	1462.88 (71.27)	1456.44 (64.82)
Grey Matter Volume (M, SD)	801.27 (47.36)	781 (45.85)	760.27 (41.12)	741.32 (40.82)
WMH volume (M, SD)	3.22 (4.42)	5.16(6.17)	5.74 (6.14)	5.49 (4.85)

Table 3.2 Two cardiometabolic condition group descriptive statistics

	Two cardiometabolic disease subgroups			
Variable	**Subgroup 5 (HT & T2D)**	**Subgroup 6 (HT & CAD)**	**Subgroup 7 (T2D & CAD)**	**Subgroup 1 No conditions**
N	133	147	8	7, 321
Age, mean (SD)	64.78(6.29)	65.93 (6.28)	66.5 (6.74)	60 (7.5)
Sex, (%F)	47 (35.34)	40 (27.21)	3 (37.5)	4, 142 (56.73)
Pack Years (M, SD)	26.1(15.64)	31.27 (20.30)	23.93 (18.63)	17.54 (13.54)
Townsend score (M, SD)	-1.45(3.12)	-1.7 (3.08)	-0.65 (4.62)	-1.97 (2.53)
Education (% degree)	58%	53%	25%	52%
BMI (Kg/m^2)	30.68(4.9)	29.62 (5.04)	31.83 (5.34)	26.05 (4.17)
Total Brain Volume (M, SD)	1461.73 (71.03)	1466.64 (71.17)	1412.98 (65.67)	1514.13 (72.25)
Grey Matter Volume (M, SD)	759.73 (47.52)	762.03 (44.97)	736.63 (38.02)	801.27 (47.36)
WMH volume (M, SD)	6.13 (7.6)	5.37 (4.53)	6.05 (4.79)	3.22 (4.42)

Table 3.3 Single cardiometabolic condition descriptive statistics

	Single cardiometabolic disease subgroups			
Variable	**Subgroup 2 (HT)**	**Subgroup 3 (T2D)**	**Subgroup 4 (CAD)**	**Subgroup 1 No conditions**
N	2, 351	115	193	7, 321
Age, mean (SD)	63.47 (6.81)	61.89 (7.57)	66.68 (5.73)	60 (7.5)
Sex, (%F)	1062 (45.17)	49 (42.61)	50 (25.91)	4, 142 (56.73)
Pack Years (M, SD)	21 (15.71)	23.74 (17.46)	23.98 (20.41)	17.54 (13.54)
Townsend score (M, SD)	-1.95 (2.63)	-1.43 (2.85)	-1.94 (2.67)	-1.97 (2.53)
Education (% degree)	55%	50%	61%	52%
BMI (Kg/m^2)	28 (4.61)	28.33 (5.45)	26.95 (4.32)	26.05 (4.17)
Total Brain Volume (M, SD)	1490.42 (69.98)	1486 (72.58)	1474.46 (67.69)	1514.13 (72.25)
Grey Matter Volume, (M, SD)	782.08 (45.76)	776.12 (48.65)	769.39 (43.56)	801.27 (47.36)
WMH volume (M, SD)	5.21 (6.25)	3.28 (3.04)	5.63 (6.31)	3.22 (4.42)

3.3.2 Associations between cardiometabolic groups and brain MRI

Table 3.4 shows associations between cardiometabolic groups and white matter hyperintensities. When fully adjusted, subgroup 2, subgroup 6 and subgroup 8 were significantly associated with WMH volume. Subgroup 8 however did not show significance when partially adjusted. When grouping by number of cardiometabolic condition, the 1 condition and 2 condition groups were significantly associated with an increase in WMH. The 3 cardiometabolic condition group was not significantly associated with an increase in WMH.

Table 3.5 shows associations with grey matter volume. When observing associations of cardiometabolic condition with volume of grey matter, all subgroups were associated with a lower volume of grey matter when compared to individuals with no conditions in subgroup 0. All subgroups were significantly associated with a lower grey matter when partially adjusted except for subgroup 4. When fully adjusted subgroups 2, 3, 5, 6 and 8 were significantly associated with a decrease, however, beta coefficients were small. For grouping of conditions, all three groups were significantly associated with a decrease in grey matter volume when both partially and fully adjusted, but beta values were also small.

In Table 3.6 for whole brain volume, subgroups 3 and 7 were significantly associated with a decrease when fully adjusted. Subgroups 3, 4 and 5 were no longer statistically significant when fully adjusted. For cardiometabolic condition groupings, groups including those with 1 condition and 2 conditions were both significantly associated with a decrease in brain volume. The group with 1 condition was no longer significantly more associated with decreased brain volume than individuals in group 0 when fully adjusted

Table 3.4 Cardiometabolic subgroups and white matter hyperintensity volume associations

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
HT (Subgroup 2)	**0.274**	**0.211 – 0.337**	**< 0.001**	**0.228**	**0.187 – 0.269**	**<0.001**
T2D (Subgroup 3)	0.073	-0.066 – 0.232	0.368	0.081	0.054 – 0.317	0.499
CAD (Subgroup 4)	**0.138**	**0.013 – 0.264**	**<0.05**	0.070	0.122 – 0.263	0.473
HT + T2D (Subgroup 5)	**0.210**	**0.013 – 0.264**	**<0.01**	0.143	0.070 – 0.357	0.189
HT + CAD (Subgroup 6)	**0.274**	**0.131 – 0.417**	**< 0.001**	**0.308**	**0.031 – 0.584**	**<0.01**
T2D + CAD (Subgroup 7)	0.377	-0.283 – 1.036	0.263	0.562	0.042 – 0.747	0.497
HT + T2D + CAD (Subgroup 8)	0.179	-0.122 – 0.481	0.243	**0.501**	**-2.18 – 1.060**	**<0.05**

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
1 Condition	**0.233**	**0.193 – 0.273**	**<0.001**	**0.263**	**0.201 – 0.324**	**<0.001**
2 Conditions	**0.247**	**0.142 – 0.352**	**<0.001**	**0.197**	**0.027 – 0.368**	**<0.01**
3 Conditions	0.179	-0.122 – 1.02	0.243	0.499	-0.0417 – 1.040	0.07

Note. Beta are standardised per SD. Partially adjusted values are for age, BMI, sex and ethnicity. Fully adjusted values are for age, BMI, sex, ethnicity, smoking, education, and Townsend. Associations significant at p=0.05 are bolded.

Table 3.5 Cardiometabolic subgroups and brain grey matter volume associations

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
HT (Subgroup 2)	**-0.003**	**-0.005** – **-0.001**	**< 0.001**	**-0.003**	**-0.006** – **0.001**	**<0.05**
T2D (Subgroup 3)	**-0.017**	**-0.025** – **-0.008**	**< 0.001**	**-0.021**	**-0.033** – **-0.008**	**<0.001**
CAD (Subgroup 4)	-0.001	-0.006 – 0.007	0.835	-0.001	-0.011 – 0.010	0.888
HT + T2D (Subgroup 5)	**-0.020**	**-0.028** – **-0.0124**	**<0.001**	**-0.013**	**-0.024** – **-0.002**	**<0.01**
HT + CAD (Subgroup 6)	**-0.010**	**-0.018 – 0.002**	**< 0.01**	**-0.021**	**-0.036** – **-0.007**	**<0.001**
T2D + CAD (Subgroup 7)	**-0.043**	**-0.073** – **-0.01**	**< 0.05**	-0.037	-0.10 – 0.027	0.255
HT + T2D + CAD (Subgroup 8)	**-0.033**	**-0.049 – 0.018**	**< 0.001**	**-0.035**	**-0.061** – **-0.009**	**<0.001**

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
1 Condition	**-0.002**	**-0.002** – **-0.001**	**<0.001**	**-0.002**	**-0.003** – **-0.001**	**<0.05**
2 Conditions	**-0.007**	**-0.01** – **-0.004**	**<0.001**	**-0.007**	**-0.011** – **-0.003**	**<0.001**
3 Conditions	**-0.014**	**-0.021** – **-0.008**	**<0.001**	**-0.015**	**-0.026** – **-0.004**	**<0.01**

Note. Beta are standardised per SD. Partially adjusted values are for age, BMI, sex, and ethnicity. Fully adjusted values are for age, BMI, sex, ethnicity, smoking, education, and Townsend. Associations significant at p=0.05 are bolded.

Table 3.6 Cardiometabolic subgroups and whole brain volume associations

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
HT (Subgroup 2)	**-0.002**	**-0.004** – **-5.754**	**<0.05**	-0.001	-0.004 – 0.002	0.686
T2D (Subgroup 3)	**-0.01**	**-0.017** – **-0.003**	**< 0.001**	**-0.013**	**-0.024** – **-0.002**	**<0.05**
CAD (Subgroup 4)	0.001	-0.006 – 0.006	0.917	0.001	-0.008 – 0.010	0.868
HT + T2D (Subgroup 5)	**-0.015**	**-0.022** – **-0.008**	**<0.001**	-0.006	-0.016 – 0.004	0.206
HT + CAD (Subgroup 6)	**-0.007**	**-0.014** – **-0.001**	**< 0.05**	-0.011	-0.024 – 0.001	0.075
T2D + CAD (Subgroup 7)	**-0.042**	**-0.069** – **-0.014**	**<0.01**	**-0.058**	**-0.114** – **-0.003**	**<0.05**
HT + T2D + CAD (Subgroup 8)	-0.011	-0.025 – 0.002	0.106	-0.021	-0.044 – 0.002	0.06

	Partially adjusted			Fully adjusted		
	Beta coefficient	95% CI	P value	Beta coefficient	95% CI	P value
1 Condition	**-0.002**	**-0.004** – **-0.001**	**<0.01**	-0.001	-0.004 – 0.002	0.446
2 Conditions	**-0.011**	**-0.016** – **-0.007**	**<0.001**	**-0.01**	**-0.017** – **-0.001**	**<0.01**
3 Conditions	-0.011	-0.025 – 0.002	0.107	-0.021	-0.044 – 0.005	0.06

Note. Beta are standardised per SD. Partially adjusted values are for age, BMI, sex, and ethnicity. Fully adjusted values are for age, BMI, sex, ethnicity, smoking, education, and Townsend. Associations significant at p=0.05 are bolded.

3.4 Discussion

3.4.1 Overview of findings

The current study investigated associations between cardiometabolic disease comorbidity and brain MRI metrics in 10, 302 UK Biobank participants. Group 1, containing individuals with no self-reported cardiometabolic condition, presented with the highest GM and whole-brain volume and lowest WMH volume in all analyses as expected. Overall, there were several significant associations between the number and type of cardiometabolic comorbidity and worse MRI measures compared to individuals with no cardiometabolic conditions, but no consistent trends. This is potentially because of the relatively small sample sizes in some subgroupings, resulting in lower reliability and unrepresentative or spurious results.

There were several instances in which subgroups showed significant associations with worse MRI measures when partially (for age, BMI, sex, and ethnicity) but not fully adjusted (for additional education, smoking and Townsend score), suggesting the role of external or environmental demographic factors, i.e., education, smoking or Townsend score, in brain-related implications of CMD. This may be particularly relevant for CAD as smoking is well established as a major risk factor for atherosclerosis which leads to CAD (Howard et al., 1998). Seven other associations in the present study similarly showed significance between cardiometabolic conditions and brain MRI when partially adjusted but were no longer associated when models were fully adjusted. This effect was highest for whole brain volume which may be due to combined effects of mechanistic pathways presenting more clearly within a global measure. One paper by Hu et al. (2020) used a machine learning model to rank sociodemographic, health and environmental factors to predict cardiovascular health in 500 US cities. When looking at HT, coronary heart disease and stroke, they found that the most common factors determined by the model as predictors of disease included cholesterol screening, obesity, leisure-time physical activity, binge drinking, and being aged 65 years or over. A review carried out by Hill-Briggs et al. (2021) similarly examined social determinants of diabetes found evidence of several social determinants that contribute to diabetes risk and outcome including economic factors, social cohesion and capital and education.

They also found that diabetes disproportionately affects racial and ethnic minority and low-income individuals in the US. The review suggested that such social determinants of health are predominantly responsible for inequalities in health outcomes, which may encompass brain health outcomes. This idea is supported by these analyses where several associations do not remain significant where environmental factors and deprivation are considered.

Backholer et al. (2016) also identified differences between men and women regarding socioeconomic inequalities as risk factors for cardiovascular conditions. The study found that while socioeconomic status and education were inversely associated with cardiovascular risk in both sexes, socioeconomic status implied an excess risk for women, whereas education did not. Suggesting mediating factors and covariates should be carefully considered when concluding the implication of cardiovascular risks on health outcomes and considered individually for each condition. The present study's findings suggest it may be essential also to consider the type of cardiovascular condition when controlling for covariates.

Furthermore, there is evidence that education can pose more significant risks for health outcomes depending on income or level of deprivation, i.e., some confounding variables may mediate others. This may be one way that covariates may influence associations in these analyses. Investigating how sociodemographic factors shape health on an individual and community level may help understand the aetiology of CMDs and multimorbidity, including factors that may differentiate those who develop CMDs in incipient stages and those who do not. Studying multimorbidity on a broader scale may help explain disparities in disease outcomes that cannot be understood through cohort studies. Differences between partially and fully adjusted models in this study indicate that such further consideration of social determinants may be helpful.

Subgroup 7 (T2D and CAD) consistently showed the largesteffect sizes for worse MRI measures. However, subgroup 7 only showed statistical significance for decreased whole brain volume. This lack of statistical significance may be explained by the low observations in this group (N=8). Therefore, due to a possible lack of statistical power, these results cannot confirm associations between subgroup 7 comorbidity and structural MRI. There are several possible

reasons for a low number of observations in this subgroup. MRI related contraindications may be one explanation as individuals with pacemakers, cardioverter-defibrillators, or prosthetic heart valves may not be able to undergo an MRI scan, creating a selection bias. There may also be fewer individuals in these subgroups because UK Biobank participants are healthier than the general population with those attending MRI scanning sessions reporting 'healthier' demographics than those who do not (Fry et al., 2017).

Before carrying out analyses, descriptive characteristics showed when visually inspected (i.e., statistical tests were not carried out) that increasing number of comorbid conditions was associated with worse MRI metrics; however, regression models do not support this after considering covariates in models. Results suggest that having two cardiometabolic comorbidities was most consistently and significantly associated with worse outcomes for MRI metrics for both partially and fully adjusted models, which was not the case for individuals with three conditions. However, the group of individuals reporting with three conditions was also underpowered (N=34) and may explain why significant associations with brain metrics were not consistently present; there is also a possibility of a survivor bias seen in the present analyses with this group.

When comparing associations found for GM, WMH and whole-brain volumes, the data showed 8 total groups out of a potential 11 (when including groupings of 1/2/3 cardiometabolic conditions) were significantly associated with a decrease in grey matter in comparison to healthy controls (no cardiometabolic conditions), whereas 5 subgroups were significantly associated with WMH and 3 with total brain volume. This is not consistent with other literature as findings of a decrease in the grey matter within cardiovascular and cardiometabolic populations are often accompanied by very similar decreases in white matter and increases in WMH (Cox et al., 2019; Roberts et al., 2014; Debette et al., 2011) with some studies emphasising progression of WMH pathology in cardiovascular diseases (Knopman et al., 2005; Moroni et al., 2018).The reason for this finding is not clear but it may be that grey matter is related to specific metabolic changes. However lack of explanation may be a result of grey matter being understudied in comparison to white matter in the context of CMD.

3.4.2 Limitations

The most significant limitation of this study was the low number of observations within some cardiometabolic subgroups, e.g., subgroup 7 (N=8) and subgroup 7 (N=34). This may reflect a selection bias or low presence of comorbidity of type 2 diabetes, hypertension and CAD (subgroup 8) and type 2 diabetes and CAD (subgroup 7). This finding may also indicate that it is more beneficial to investigate comorbidity by the number of conditions rather than splitting into types of condition, as the latter may not provide more specific information regarding outcome effects of cardiometabolic disease; however, this is unlikely.

Another limitation of this work is that alternative risk factors such as diet, psychosocial factors or physical activity have not been considered, although strongly implicated in developing cardiometabolic disease. One meta-analysis studying epidemiological links between mental health and cardiometabolic disease highlights bidirectional links between depression and cardiovascular disease (Penninix, 2016). Another study showed that 9% of CAD patients presented with generalised anxiety disorder compared to 5% of healthy controls (Tušek-Bunc & Petek, 2016). Physical activity is another significant factor that plays a role in health related outcomes of CMD and was not considered here, and may explain some unexpected results.

There is some evidence to suggest that vascular risk factors are associated with worse brain MRI outcomes later in life when exposure is during midlife (Debette et al., 2019). This study, however, was cross-sectional, and both cardiometabolic and MRI data were taken at the same data instance. This makes reverse causation a possibility within this data, for which it is not possible to parse out which came first, the cardiometabolic condition or worse brain health.

The generalisability of these results is limited by selection bias within the UK Biobank in which one study from by the Department of Population Health at the University of Oxford (Fry et al., 2017) found that UK Biobank participants were less likely to be obese, smoke, be socioeconomically deprived and have fewer self-reported health conditions when compared to the general population. 95.6% of the UK Biobank cohort are of white ethnicity, which is limiting when considering the clinical applications of cardiometabolic research findings as

there are significant differences in the presentation of cardiovascular disease for different populations. For example, classical risk factors such as waist circumference or BMI cut-offs vary between populations (Zhu et al., 2005; Chaturvedi, 2003). Finally, the methodological definitions of cardiometabolic diseases used for these analyses were self-reported and may not be as accurate as physical measures such as blood pressure, pulse rate or blood glucose.

3.4.3 Future work

One area of future work includes replicating associations with larger groups of individuals. The current study analysed single data collection instances; therefore, future studies could use longitudinal data such as future follow-up waves in UK Biobank to investigate the association of cardiometabolic disease with brain health over time, with the ability to conclude directionality.

Brain MRI metrics in this study were limited. Future investigations may find more specific findings when looking closer at metrics of neurodegeneration such as white matter tract integrity or additional structural measures, particularly as changes are likely to be smaller in midlife. This may include MRI metrics such as fractional anisotropy, mean diffusivity, or hippocampal volume.

This study has also highlighted the need for future studies to look more closely at lifestyle factors and consider these in more detail. It is still unclear how lifestyle factors such as socioeconomic inequalities or psychosocial factors may influence and contribute to brain-related health outcomes of cardiometabolic conditions.

Recent Genome-Wide Association Studies (GWAS) have found that, like most complex diseases, there is heritability of HT, CAD and T2D. Common variants have been studied in the literature, many of which have been implicated in obesity and insulin resistance (Cheung & Li., 2012). One study found that offspring of hypertensive individuals presented with heritable abnormal glucose metabolism (Friedman et al., 2005), suggesting overlapping genetic predispositions for comorbidity where consideration of genetics may be helpful when predicting potential effects of comorbidity.

3.4.4 Conclusion

Overall, the associations between cardiometabolic comorbidity and brain MRI provide preliminary evidence for differential effects of cardiometabolic conditions on brain health outcomes. Most associations were seen for grey matter volume in comparison to whole brain volume and WMH. However, it is essential to note that findings of cardiometabolic subgroupings and worse brain health did not present any clear trends, and there were no consistent findings between groups. Despite the lack of evidence for clear associations between type of cardiometabolic comorbidity and brain MRI, to our knowledge, this is the first study to investigate potential differences between different patterns of comorbid cardiometabolic conditions concerning brain health.

Chapter 4: Elevated Lipoprotein A and brain MRI phenotypes

4.1 Introduction

4.1.1 Lipoprotein A and Alzheimer's disease

The role of lipids is vital to normal brain functioning, e.g., blood-brain barrier (BBB) function, myelination, inflammation, APP processing, receptor signalling and energy metabolism (Chew et al., 2020) and recent studies looking at differences between healthy cognitive functioning vs abnormal brain-related pathology have stressed the importance of lipid disruption (Schilling et al., 2014; Proitsi et al., 2017). Studies investigating disturbed lipid functions in the context of brain pathology and histology have found that a disrupted brain lipidome can lead to abnormalities in the functioning of astrocytes and microglia. One study showed that unbalanced lipid homeostasis affected signal transduction via membrane phospholipids (Farooqui et al., 1988). GWAS have also implicated lipids and lipid metabolism in Late-onset Alzheimer's Dementia (AD) pathology (Kunkle et al., 2019) with lipidomic analyses of different brain areas confirming age-related lipid alterations in disease states (Proitsi et al., 2017) and one study found that the use of statins reduced the risk of incident AD (Sparks et al., 2008). Ferguson et al. (2020) found that when looking at n=395,769 healthy individuals in the UK Biobank, individuals who were at genetic risk for AD presented with different biomarker and lipid profiles to those who were not at genetic risk, supporting the use of lipid assessment within healthy populations to indicate who is at a higher risk of cognitive decline. There is growing evidence to implicate lipid dysfunction in abnormal brain ageing, creating the potential for lipid biomarkers.

Lipoprotein A (LpA) is a low-density lipoprotein variant that is considered to be a reliable CVD risk factor. There is evidence of causal associations where observational and genetic evidence strongly support a causal relationship between high plasma concentrations of LpA and increased risk of CVD–related events, such as myocardial infarction and stroke (Larrson et al., 2020). The Copenhagen City Heart Study found that high LpA levels were predictive of a 3-to-4-fold increased risk of myocardial infarction in the general population (Kamstrup et al., 2008). It has also been suggested that increased LpA serum concentrations may play a role in determining clinical AD outcomes via increased risk for cerebrovascular disease (Iturria-Medina, Hachinski and Evans, 2017).

However, there is little data in comparison investigating associations of LpA with subsequent dementia and impairment in cognition. Previous studies have reported observational associations with LpA and AD, but the evidence is inconclusive. Solfrizzi et al. (2002) found that LpA serum concentrations were significantly associated with an increased risk for AD, independently of *APOE* genotypes and sex. Emanuele et al. (2004) studied LpA isoforms in 73 AD patients matched by age and gender with healthy controls. They found that isoforms that determine higher levels of LpA were associated with the age of AD onset but not progression to AD. Iwamoto et al. (2004) found LpA levels were highest in vascular dementia patients (n=46), but with LpA levels still higher in the AD group (n=150) than in healthy controls (n=150), suggesting LpA may contribute to the development of AD through vascular pathways. However, papers present conflicting evidence; Kunutsor et al. (2016) found in a prospective cohort study of 2,532 Finnish men that LpA was inversely associated with future risk of dementia. This finding may be due to a genuine inverse relationship or lack of association, although it has also been proposed that the lack of association may be confounded by competing risks and survival bias in which individuals with high LpA and CVD do not live long enough to progress to AD.

4.1.2 Lipid biomarkers and brain MRI phenotypes

Studies investigating the nature of alterations in brain tissue in AD individuals show white matter hyperintensities (WMH) are a prominent feature of both CVD and AD. Although WMHs are associated with both CVD and AD, it is not clear how to differentiate vascular and AD pathophysiological changes and contributions to brain damage or cognitive decline. One multicentre study of WMHs in n=2,699 stroke patients found that cholesterol was a more significant risk factor for WMH when individuals were hypertensive than in non-hypertensive individuals where age was a more significant risk factor (Ryu et al., 2017). These findings suggest that lipid profiles may contribute to the development of WMH to a varying extent depending on whether other relevant risk factors are present, highlighting the importance of studying such biomarkers profiles in relation to other risk and lifestyle factors.

Some studies suggest that development rather than presence of WMH is associated with adverse outcomes such as cognitive decline, as more rapid deterioration is likely to disrupt neural networks (Filippi et al., 2020). A twin study by Sachdev et al. (2016) found that heritability for WMH volume was between 64% to 77% depending on the brain region, suggesting genetic factors may also play a role in developing WMHs, it is possible that such variants are pleiotropic and may be involved in lipid function. Some neuroimaging methods studying structural changes occurring to white matter integrity have attempted to quantify more minor changes to white matter, which is relevant to early disease pathology, including investigating microstructure. One study by Cox et al. (2016) analysed white matter in 3,513 UK Biobank individuals with MRI data and associations with age to understand how white matter microstructure varies over the lifespan. The study investigated multiple microstructural characteristics, including fractional anisotropy (FA) and mean diffusivity (MD). Additional analyses were carried out for FA and MD in which 22 significant white matter tracts were identified; these were used within this analysis and discussed in the methodology section below.

4.1.3 Genetic correlates of LpA

The use of genomics in cardiovascular medicine is increasing, and advancements in cardiogenomics are helping to move towards a precision medicine approach to preventing cardiovascular diseases in individuals at risk. The American Heart Association (AHA) suggested 30 “medically actionable genes” related to CVD, based on a paper published by the American College of Medical Genetics and Genomics, including TTR and PCSK9, both of which play critical roles in lipid homeostasis (Kalia et al., 2016). These studies show that identifying genetic variations that indicate the heritability of cardiovascular conditions and pathology is helpful for both disease prevention and gaining a closer understanding the role of lipids. One of the most common cardiovascular conditions is atherosclerosis, which manifests clinically as coronary heart disease and stroke. LpA is an established CVD risk factor and plays a role in the development of atherosclerosis.

Heritability of elevated levels of LpA is thought to be around 80 to 90%, where the LPA gene locus explains a substantial proportion of the variance (Welsh et

al., 2020). One study reports around 30 – 70% of LpA heritability to be accountable by the copy number variant (CNV) Kringle-IV (KIV)-2 in the LPA gene, with low numbers of KIV repeats resulting in higher levels of LpA concentrations (Kronenberg., 2016). Other SNPs known to contribute to LpA levels have also been identified. One meta-analysis of five LpA-GWAS reported a total of 31 SNPs found to be significantly associated with LpA concentrations in both LPA and *APOE* regions (n=13,781) (Mack et al., 2017). Inheritance of the *APOE* e4 allele may be of relevance as it is involved in regulating lipid metabolism, and is associated with several cardiovascular factors, including CVD, LDL, total cholesterol, and cerebrovascular related cognitive decline (Rojas et al., 2018).

4.1.4 Study rationale

LpA has been used as a reliable risk factor for CVD and plays a role in developing AD pathologies. Studies investigating causal roles of LpA have found that there does appear to be a causal relationship between elevated LpA concentrations and subsequent CVD. However, beyond the ε4 allele of apolipoprotein E (*APOE*), comparatively little is known about whether genetic variants associated with a higher risk of cardiovascular disease can be considered risk genes for AD through potential genetic pleiotropy (Lippi et al., 2019). Previous studies investigating genetic links between CVD and AD have found significant overlap in risk factors; however, it has been challenging to draw conclusions about the role of LpA in the development of neurodegeneration due to lack of evidence. This study examines whether individuals with elevated blood LpA and genetically elevated LpA are associated with brain phenotypes of dementia, including white matter structure, brain volumes and cognitive performance in the UK Biobank.

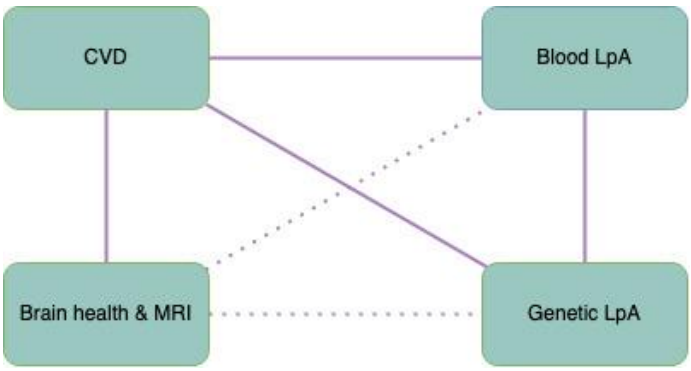

Figure 4.1 Strength of evidence between LpA, CVD, brain and cognition

4.2 Methods

4.2.1 Biomarker collection and processing

LpA biomarker levels were analysed in UK Biobank from blood samples at baseline from UK Biobank participants. This study investigated baseline measure of lipoprotein A which was reported in units of measurement of nmol/L. Values more than 5 SDs from the mean were removed (M=44.64, SD=49.21).

This document outlines quality control approaches that were undertaken to mitigate random errors and bias in biomarker samples: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf.

4.2.2 LpA GWAS variants

The variants used to create the PGR score of blood LpA were taken from Pan et al. (2019), where they identified 9 SNPs associated with LpA concentrations as instrumental variables for an AD Mendelian Randomization study. To identify these variants, Pan and colleagues obtained summary-level data from Clarke et al. (2009)'s investigation of SNPs from the LpA region for an association analysis on level of LpA. Pan et al., (2019) then selected SNPs that achieved significance at $P<1.0\times10-6$ from Clarke et al. (2009). Data carried forward from both studies for the present analysis was found in individuals with European ancestry only.

4.2.3 LpA genetic score

When creating the PGR score for elevated LpA, genetic QC excluded individuals with non-white British ancestry, self-report vs genetic sex mismatch, putative sex chromosomal aneuploidy, excess heterozygosity, missingness rate > 0.1. Additionally, those who reported a neurological condition at baseline were excluded. All remaining individuals were included in the analysis in which PGR scores were calculated. LDPred was used to create genetic risk scores in which 9 LpA-raising alleles inherited at each variant were included in the score, weighted by each variant's association with absolute change in LpA mass concentration, which was taken from the effect sizes given from the GWAS. Further details about genotyping methods used in UK Biobank are outlined in chapter 2: Methodology.

4.2.4 MRI and cognitive measures

Brain MRI phenotypes included in these analyses were volumes of grey matter (log), white matter (log), white matter hyperintensity (log), whole brain (log), left hippocampus and right hippocampus. Segmented regions of each left and right hippocampus, including hippocampal head, tail and body, were included. Right and left grey hippocampal matter have been included as outcome variables for the first time in this thesis to investigate grey matter more closely. This is different to left and right hippocampus volumes that have been included in subsequent chapters that refer to white + grey matter. We also included general factors of white matter microstructure as a latent variable. We constructed general factors of FA (gFA) and MD (gMD) using principal components analysis based on 22 tracts (Cox et al. 2016) where gFA eigenvalue = 10.98, r2 = 49.89% and gMD eigenvalue = 11.79, r2 = 53.57%. The 22 tracts included left and right: acoustic radiation, anterior thalamic radiation, cingulate gyrus, corticospinal tract, inferior fronto-occipital fasciculus, inferior longitudinal fasciculus, posterior thalamic radiation, superior longitudinal fasciculus, superior thalamic radiation, uncinate fasciculus, and both the forceps major and the forceps minor. Handedness was not included as an exclusion criterion in these analyses; a study by Cox et al (2016) showed that handedness does not have an effect on FA and MD in a meaningful way.

Cognitive measures included in analyses included: fluid intelligence, prospective memory, numerical memory, reaction time, total trail making (A+B), and symbol digit substitution. More details about how these variables were derived are in Chapter 2: Methodology.

4.2.5 Covariates

All association models were run twice: once partially adjusted and then fully adjusted. Partially adjusted models included age, sex, BMI, PCs 1-8 for population stratification, genotyping chip and fully adjusted for additional, *APOE* e4 dose, cardiovascular medication, smoking (never vs ever), education and Townsend social deprivation score.

4.2.6 Statistical analyses

Regression models were run with 16 brain MRI and 6 cognitive phenotypes to investigate associations between both genetic risk scores for LpA and blood LpA, in which covariates were considered. Observational correlations for covariates and PGR-LpA and blood LpA were also reported

4.3 Results

4.3.1 Demographic statistics

Table 4.1 shows demographic statistics of individuals split by male and female sex. Table 4.2 shows summary statistics used to create PGR scores of elevated LpA taken from Pan et al. 2019; Causal Effect of Lp(a) Level on Ischemic Stroke and Alzheimer's Disease. Table 4.3 shows correlations between blood and PGR of elevated LpA used in this chapter.

Table 4.1 Demographic statistics by male and female sex

Male vs Female	N	Age (years)	BMI (kg/m2)	Smoking (% ever)	Townsend deprivation score	LpA (nmol/L)
Female	21,005	63 (7.41)	26.1 (4.75)	33.65%	-1.86 (2.72)	1.65 (0.26)
Male	18,690	64.32 (7.65)	26.46 (3.91)	40.58%	-1.94 (2.71)	1.44 (0.22)

Note: LpA (nmol/L) is mean (SD)

Table 4.2 GWAS significant SNPs associated with *LpA*

SNP	Chr	Locus	EA/OA	EAF	Beta	SE	p
rs10455872	6	*LPA*	G/A	0.07	1.18	0.04	3.6×10^{-166}
rs3798220	6	*LPA*	C/T	0.02	1.27	0.08	5.9×10^{-51}
rs11751605	6	*LPA*	C/T	0.16	0.50	0.04	5.9×10^{-28}
rs10945682	6	*LPA*	G/A	0.64	0.32	0.04	1.8×10^{-17}
rs6919346	6	*LPA*	C/T	0.83	0.43	0.05	1.6×10−16
rs3127596	6	*LPA*	G/A	0.30	0.30	0.04	1.5×10−14
rs10755578	6	*LPA*	G/C	0.48	0.27	0.04	3.4×10−13
rs3798221	6	*LPA*	G/T	0.81	0.28	0.05	2.0×10−9
rs6415084	6	*LPA*	T/C	0.49	0.22	0.04	2.7×10−9

Note. Chr = chromosome; EA = effect allele; OA = other allele; EAF = effect allele frequency; SE = standard error.

Table 4.3 Correlations between covariates and blood and PGR of elevated LpA

	PGR LpA	LpA blood
Age	<-0.012 (0.59)	**0.14 (0.017)**
Sex	-0.07 (0.04)	0.12 (0.06)
BMI (kg/m2)	<-0.012 (0.83)	0.15 (0.19)
Smoking (never vs ever)	-0.15 (0.069)	-0.22 (0.078)
PGR LpA	-	**0.56 (<0.003)**

Note: Values are: correlation coefficient (p-value)

4.3.2 Associations between LpA and brain MRI

Table 4.4. shows that 9 brain MRI measures were significantly associated with blood LpA when partially adjusted, these include WMH (ß = 0.04, p = <0.001), gFA (ß = 0.02, p = 0.006), gMD (ß = 0.01, p = 0.04), left hippocampus (ß = 7.53, p = 0.005), right hippocampus (ß = 8.46, p = 0.002), left hippocampal body (ß = ,-1.96 p = 0.01) and tail (ß = -1.09, p = 0.05), and grey matter of both left (ß = 5.99, p = 0.01) and right hippocampus (ß = 4.98, p = 0.04). When these models were further adjusted for CVD medication, BMI and *APOE* dose the associations that remained significant were for gMD (ß = 0.03, p = 0.005) and left hippocampal body (ß = -3.34, p = 0.04). Additionally, whole brain volume (ß = -2.38, p = <0.001) and grey matter volume (ß = -3.12, p = <0.001) were associated with blood LpA for fully adjusted models only. When looking at all significant associations within fully adjusted models, they are in the deleterious direction, e.g., lower whole brain volume and higher mean diffusivity associated with higher levels of LpA. However, within partially adjusted models, left and right hippocampal volumes and left and right grey hippocampal volumes are positively associated, suggesting higher brain volumes with increased concentration of blood LpA.

Table 4.5 shows that no brain MRI measures were significantly associated with PGR of elevated LpA when partially adjusted. General factor of mean diffusivity (β = 0.02, p = 0.03) and general factor of fractional anisotropy (β = -0.01, p = 0.02) were associated with PGR of elevated LpA in fully adjusted models. No other MRI variables were associated with PGR of elevated LpA. Table 4.6 shows fully adjusted associations between cognitive measures and blood LpA, and PGR of elevated LpA. The table shows that no cognitive measures were associated with blood or genetically elevated LpA measures in these analyses. Effect sizes were small, and no associations were close to significance

Table 4.4 Associations between blood LpA and structural brain MRI

	Partially adjusted		Fully adjusted	
	β	p	β	p
Whole brain volume (log)	-3.99	0.09	**-2.38**	**<0.001**
Grey matter volume (log)	-4.71	0.08	**-3.12**	**<0.001**
White matter volume (log)	-3.03	0.36	-1.15	0.08
WMH (log)	**0.04**	**<0.001**	0.01	0.60
gFA	**0.02**	**0.006**	-0.01	0.08
gMD	**0.01**	**0.04**	**0.03**	**0.005**
Left hippocampus (mm3)	**7.53**	**0.005**	-0.84	0.89
Right hippocampus (mm3)	**8.46**	**0.002**	2.74	0.65
Right hippocampal body (mm3)	-1.37	0.08	-2.62	0.12
Right hippocampal head (mm3)	-0.76	0.51	-3.35	0.17
Right hippocampal tail (mm3)	-0.50	0.21	-0.76	0.37
Left hippocampal body (mm3)	**-1.96**	**0.01**	**-3.34**	**0.04**
Left hippocampal head (mm3)	-0.91	0.42	-3.97	0.10
Left hippocampal tail (mm3)	**-1.09**	**0.05**	-2.01	0.09
Grey matter left hippocampus (mm3)	**5.99**	**0.01**	3.06	0.53
Grey matter right hippocampus (mm3)	**4.98**	**0.04**	1.12	0.82

Note. Bold = significant p<.05. Models are partially adjusted for age, BMI, sex, genotyping chip, 8 genetic principal components and fully adjusted for additional cardiovascular medication, *APOE* e4 dose, smoking, education and social deprivation.

Table 4.5 Associations between PGR of elevated LpA and MRI measures

	Partially adjusted		Fully adjusted	
	β	p	β	p
Whole brain volume (log)	5.94	0.69	4.53	0.86
Grey matter volume (log)	-5.96	0.97	3.11	0.46
White matter volume (log)	1.78	0.52	4.62	0.36
WMH (log)	0.01	0.71	0.01	0.39
gFA	0.01	0.71	**-0.01**	**0.02**
gMD	0.02	0.63	**0.02**	**0.03**
Left hippocampus (mm3)	0.96	0.19	1.21	0.59
Right hippocampus (mm3)	-0.68	0.77	2.01	0.66
Right hippocampal body (mm3)	1.12	0.09	1.64	0.21
Right hippocampal head (mm3)	1.31	1.18	2.78	0.14
Right hippocampal tail (mm3)	0.23	0.51	0.67	0.31
Left hippocampal body (mm3)	0.91	0.17	1.56	0.22
Left hippocampal head (mm3)	1.33	0.16	3.48	0.06
Left hippocampal tail (mm3)	0.08	0.85	1.22	0.17
Grey matter left hippocampus (mm3)	0.23	0.91	3.82	0.31
Grey matter right hippocampus (mm3)	0.83	0.68	4.51	0.24

Note. Bold = significant p<.05. Models are partially adjusted for age, BMI, sex, genotyping chip, 8 genetic principal components and fully adjusted for additional cardiovascular medication, *APOE* e4 dose, smoking, education, and social deprivation.

Table 4.6 Associations between cognition and blood LpA, and PGR of elevated LpA

	LpA blood		LpA PGR	
	β	p	β	p
Fluid intelligence	<0.001	0.32	0.006	0.74
Prospective memory	-0.01	0.26	-0.002	0.67
Numerical memory	<0.001	0.66	-0.032	0.15
Reaction time	0.02	0.51	0.66	0.57
Trail making	<0.001	0.19	<0.001	0.89
Symbol digit	<0.001	0.75	-0.05	0.48

4.4 Discussion

This study investigated associations between elevated blood LpA and genetically elevated LpA and brain MRI in 39, 695 individuals. We replicated previous associations with blood lipid LpA and found that blood LpA was associated with whole-brain volume, grey matter volume, gMD and left hippocampal body volume. PGR of elevated LpA only showed associations with gMD and gFA. Although findings between PGR and blood LpA did not overlap exactly, overall findings do suggest both blood and genetically elevated LpA were associated with higher MD. A higher value of MD indicates an increase in overall diffusion; this may be due to poorer WM microstructural integrity or high-water content due to inflammation or oedema, suggesting higher levels of elevated LpA may be associated with this endophenotype. No cognitive measures were associated with blood or PGR of elevated LpA, suggesting no differences in cognition were evident amongst those with the highest levels of LpA and those with the lowest levels of LpA.

4.4.1 PGR LpA vs blood LpA

When comparing blood and PGR associations with brain MRI phenotypes, 9 out of 16 showed conflicting directions, in which several PGR LpA associations showed unexpected directionality, i.e., higher brain volumes. There could be several reasons that blood LpA and genetic risk of elevated LpA showed conflicting directionality. Emanuele et al. (2004) compared APOA isoforms to LpA levels in the blood and found no differences between isoform expressions and LpA levels in AD patients and controls, even when controlling for *APOE*. However, when comparing the age of onset for those with AD, individuals with at least one APOA isoform had a significantly higher mean age of onset than those who did not have an APOA isoform (76 vs 66, p=0.01). This suggests that genetic determinants of LpA concentrations may not present significantly different LpA levels as measured by circulating blood samples but may still influence pathogenesis. High blood LpA and genetically elevated LpA may reflect different pathogenesis that may ultimately influence AD progression in separate ways, e.g., blood LpA may be more affected by cardiovascular health or lifestyle factors. This potential difference in pathways may also explain why some blood LpA associations with

brain volumes no longer remained when models were fully adjusted, but this effect was not seen for genetically determined LpA.

Furthermore, the cross-sectional nature of these analyses could be one reason some LpA associations seemingly show unexpected directionality in association with MRI measures. The age-of-onset may be a crucial factor to consider when investigating the risk of AD and demonstrates that cross-sectional studies cannot draw conclusions regarding the role of risk factors over time or may portray an incomplete picture depending on when data are taken.

Another explanation for conflicting PGR and blood directionality may be that the genetic risk model in these analyses may not have been a good predictor for levels of LpA in midlife; during the discovery of the genetic variants, selection may have been biased for extreme values of high LpA, or in those most at risk for elevated LpA and therefore not representative of LpA levels within a general population. The clinical utility of this genetic risk score based on summary data may not be applicable in wider populations, such as the UK Biobank. It is possible that the ancestral backgrounds of the summary statistics were not compatible with the LpA variants present in the UK Biobank study population (Clarke et al., 2009; Pan et al., 2019). Participant demographic and clinical characteristics may have differed to the extent that the risk score was not generalisable to individuals considered healthier than the general population, particularly as LpA concentrations are influenced by lifestyle and demographic factors (Wilson et al., 2019).

4.4.2 Sex differences

No sex differences were observed in additional analyses. Currently, the literature is conflicting and inconclusive regarding whether LpA levels vary significantly amongst men and women (Welsh et al., 2020; Wilson et al., 2019); it is known that cardiovascular risk factors are more prevalent in men than in women under the age of 60, many of which are known to be associated with the development of AD (Rojas et al., 2018). There is evidence that sex differences could be an essential factor for the stratification of individuals. Ferguson et al. (2020) reported that out of 33 biomarkers, 16 showed sex vs *APOE* genotype interaction, including ApoA, ApoB and cholesterol. Alternatively, Gong et al.

(2021) examined whether sex differences in cardiovascular risk factors were related to all-cause dementia in women within the UK Biobank cohort and found that several cardiovascular factors were associated with dementia in women but not for men. It may be the case that the UK Biobank sample may be underpowered to detect minor effects as individuals with cardiovascular conditions appear underrepresented in this cohort (Fry et al., 2017).

4.4.3 *APOE* e4

APOE e4 interactions may be an area of future study, which might help to explain why the directionality of some results in partially adjusted models differed from fully adjusted models. Mooser et al. (2000) found that while elevated LpA levels were a risk factor for AD in those who carry *APOE* e4 genotype, non-e4 carriers' LpA levels were associated with reduced risk of AD in those over 80 years of age. While the role of *APOE* genotype was not examined in this study, future studies investigating this association may aid in determining the role LpA may play in the development of AD.

4.4.4 Limitations

There is evidence that variations in diet, circadian rhythmicity, seasonal cycles, medical conditions may all introduce biases and were not considered within these analyses (Henriksen et al., 2014). When using known CVD markers such as LpA, it is essential to consider multimorbidity and incident diseases in studying biomarker profiles; for example, inflammatory processes are likely to be more common with multimorbidity and within elderly populations (Henriksen et al., 2014). Studying biomarkers in isolation makes it difficult to characterise and recognise when it is indicative of a disease state. Particularly when using genetic instruments, as further unmeasured confounders can occur between genetic variants and outcomes. For example, pleiotropic effects of variants within the LpA region influencing outcomes of interest through other pathways such as HDL. Additionally, confounders that have been considered within these analyses may not perfectly characterise them and could result in subsequent residual confounding such as CVD medication.

Additionally, there is conflicting evidence of what is considered high levels of circulating LpA and what cut off points should be to investigate the role of LpA. One cohort study by Kamstrup et al. (2009) studied how genetically elevated levels of LpA increased the risk of myocardial infarction in which cut-off for high levels of LpA was at 30 mg/dL (67th percentile) and 85 mg/dL and 120 mg/dL, were approximately 90th and 95th percentiles. However, the LpA HORIZON trial uses 70mg/dl as an inclusion criterion at initial screening, and the Heart UK consensus statement suggested that LpA levels between 77mg/dL and 150 mg/dL should be considered high (Jaimini et al., 2019). This lack of standardisation for high LpA levels is partly due to epidemiological differences and differences amongst individuals in different ethnic groups (Wilson et al., 2019).

4.4.5 Conclusion

This is the first study to study and compare the associations of blood and PGR of elevated LpA and brain MRI. We found that both blood LpA and PGR of elevated LpA was associated with worse brain MRI phenotypes, particularly with higher gMD. This analysis provides some initial evidence that both blood and genetically elevated metrics of LpA have potential to be valuable in examining differences in brain health; however, we also suggest that there are discrepancies between the two measurements which should be further investigated.

Chapter 5: Polygenic risk of Alzheimer's Dementia and brain MRI and cognition

A version of this chapter is published in Neuropsychopharmacology:

Tank, R., Ward, J., Flegal, K.E. et al. Association between polygenic risk for Alzheimer's disease, brain structure and cognitive abilities in UK Biobank. Neuropsychopharmacology. 47, 564–569 (2021). https://doi.org/10.1038/s41386-021-01190-4

5.1 Introduction

According to the WHO, dementias affect about 47.5 million people worldwide, 60-70% of which are cases of Late-Onset Alzheimer's disease (AD) (2016) and estimates predict that 135 million people will live with dementia in 2050 (WHO, 2020). As a major health problem of the 21st century, understanding prognosis and improving diagnosis for AD is necessary. However, AD is a progressive disease with insidious onset, and prior to clinical diagnosis and subsequent progression, individuals will already have experienced considerable cognitive deficits and attendant brain pathology (Counts et al., 2017).

5.1.1 Definition and characterisation of AD

For individuals with AD, functional impairments are typically due to the deterioration in cognitive and behavioural domains such as attention, executive functioning, and memory (Garcia-Ptacek et al. 2016; Jungsu, Basak & Holtzman, 2009; Van der Elst) in which individuals tend to experience difficulties with language, short-term memory loss, mood swings and presentation of challenging behaviour (Hedden & Gabrieli, 2005). Although a definitive Alzheimer's diagnosis can only be made post-mortem via brain autopsy, a probable diagnosis can be made when symptoms interfere significantly with an individual's daily functioning. A probable diagnosis is based on neuropsychological assessments that evaluate mental and cognitive status. Medical history, blood tests, and brain imaging may be used to rule out other causes for symptoms. Neuropsychological criteria for diagnosis include objective evidence of memory deficits, executive functioning, disturbances in consciousness, or neurological pathology (McKhann et al., 2011).

AD progresses over many years and stages preceding clinical AD include preclinical and prodromal dementia. According to NIA-AA criteria, preclinical AD refers to a stage in which cognitively normal individuals present Alzheimer's pathology, whereas prodromal stages refer to individuals experiencing objectifiable cognitive impairments. Predementia stages include subjective cognitive impairment (SCI) and mild cognitive impairment (MCI). When memory loss is a predominant symptom of cognitive impairment, around 15% of individuals with amnestic MCI develop probable AD per year (Grundman et al.,

2004). Memory related decline in such instances has been related to changes in the entorhinal cortex and medial temporal lobe due to pathologies such as beta amyloid-beta peptide accumulation and neurofibrillary tangles; for this reason, the entorhinal cortex and hippocampus have been areas of interest in AD research. MCI patients do not always progress to Alzheimer's as it is possible to stay stable or revert to normal cognitive functioning. However, this pathological process begins decades before the onset of AD symptoms. For this reason, much research has aimed to identify methods of risk prediction and opportunities for early intervention.

5.1.2 Genetic correlates of AD

The most considerable risk factor for AD after age is the ε4 allele of the apolipoprotein E gene (*APOE*). (Kim, Basak & Holtzman, 2009). *APOE*, a gene on chromosome 19, codes for a protein that plays a role in the metabolism of fats in the body. *APOE* has been associated with cardiovascular disease and Alzheimer's disease; it typically plays a role in transporting lipoproteins and cholesterols and the metabolism of Aß. The *APOE* protein enhances the breakdown of Aß; however, the ε4 variant is not effective at enhancing the breakdown and metabolism of Aß, creating a pathophysiological vulnerability for AD (Kim, Basak & Holtzman, 2009). The ε4 allele is associated with the most risk for AD, and the ε2 allele is protective. The approximate prevalence for Caucasians in the US for the six permutations of *APOE* variants is as follows ε2/2 (1%), ε2/3 (22%), ε2/4 (2%), ε3/3 (58%), ε3/4 (14%) and ε4/4 (3%). At least one copy of the ε4 allele is found in around 65% of patients with AD and around 10 – 15% of the general population (Zhao et al., 2020). The presence of the ε4 allele has also been associated with accelerating Aß deposition and has been implicated in the regulation of tau phosphorylation (Hedden and Gabrieli, 2005). One study also found that ε4 carrying neonates had lower brain volumes in parietal and occipital lobes compared to non-carriers, suggesting strong evidence for predisposing genetic factors (Ferreira et al., 2017). However, it is essential to remember that presence of the *APOE* ε4 allele is not necessary or sufficient for AD to develop.

In addition to *APOE* ε4, there is increasing interest in other genetic contributions to AD, and other genetic risk factors are also known to play a role in developing

AD through pathways such as inflammation or mediating amyloid and tau pathology; these genetic factors can take the form of risk genes. Heritability estimates range from 50% to 79%, with both common and rare variants contributing to risk and discovery of relevant genetic variants for AD has increased substantially in recent years. One study investigated differences between PGR of AD for ε4 carriers and non-carriers. They found that when comparing ε4 carriers with non-carriers of the same PGR, carriers had a significantly earlier age of onset. The median survival age for ε4 carriers was 75; this number changed to 73 for individuals with high PGR or 80 for those with low PGR (Shi et al., 2019). The findings of this study, along with many others, suggest that identifying additional genetic risks for AD could provide a more robust risk prediction, and better identification of high-risk individuals (Dudbridge et al., 2013; Lupton et al., 2016; Stocker et al., 2018; Shi et al., 2019).

The International Genomics of Alzheimer's Project (IGAP) recently released results and summary statistics from a meta-analysis of gene candidate studies for AD. This data file included ~7 million genotyped or imputed SNPs, with 8, 527 cases of AD and 11, 312 controls (Kunkle et al., 2019). The study showed at least 20 loci associated with AD susceptibility at the genome-wide significance level ($P < 5 \times 10\text{-}8$). The IGAP consortia findings have significantly contributed to advancing genetic research for AD since they were published in 2013. However, it is unclear to what extent individual genetic variants can predict risk and conversion to AD. One review paper by Stocker et al. (2018) found that out of 18 PGR studies for AD, all were found to be predictive of AD status or conversion to AD. Investigating these additional genetic risks may reveal more about the pathology and cognitive deficits experienced leading up to AD, which has not been widely studied.

Genetic variants have also been associated with known risk factors for AD. For example, the immune response is highly active in AD, and some research groups have been studying the contribution of genetic risk from immune-related genes (Mrdjen et al., 2019). However, neurogenetic regulatory mechanisms are very interactive, and it is becoming more common to use methods capable of considering multiple complex genetic phenotypes. With the identification of

relevant genetic variants becoming more robust due to improved GWAS methods and data resources, PGR scores are also increasing.

5.1.3 Additional risk factors

Vascular problems are the most common comorbidity alongside AD, and for decades there has been a clear association with cardiovascular abnormalities and development of AD; however, the nature of the relationship is not clear (Van Dijk et al., 2004). This may be partly due to reduced blood flow and oxygen to the brain, as well as a breakdown of the blood-brain barrier, which regulates necessary metabolites such as clearing of toxic beta-amyloid and tau proteins (Van De Haar et al., 2017). Cardiovascular risk factors, including hypertension, stroke, diabetes, and cholesterol, can lead to inflammation or cerebrovascular disease (CVD) in the brain, markers of CVD including cortical atrophy, white matter hyperintensities known as microvascular changes, cortical and subcortical infarcts (Tosto et al., 2015). There is evidence from the Vanderbilt Memory and Ageing Project that low cardiac index is associated with a reduction of blood flow to the temporal lobe but not to other brain areas (Jefferson et al., 2017). This may partly explain why low cardiovascular health is a risk factor and common comorbidity of AD. Other studies suggest this brain health to heart health connection may be due to more epidemiological links such as obesity, sex or age.

The associations between age, sex, and *APOE* have been termed the “triad risk of AD”, each a well-established risk factor for AD, and recent evidence suggests complex interplay occurs. It is understood that the female sex is a risk factor for developing AD, with around 60% of cases. However, this is not explained by greater longevity (Brinton et al., 2015). There has been research investigating the role of menopause in AD showing estrogen plays a role in cardiovascular health and normal cognitive functioning, in which perimenopausal women can experience a decline. However, evidence from longitudinal cohort studies shows that cognitive decline seen in perimenopausal women can be limited, and individuals showed improvements after 18 months (Morgan et al., 2018). This menopause-related cognitive impairment may be due to estrogen-regulated systems such as sleep, circadian rhythms and specific cognitive domains that act as a risk factor for AD development. However, there is little research to

differentiate MCI from neurological symptoms of perimenopause and some people are suspected to be misdiagnosed with MCI.

5.1.4 Brain MRI phenotypes in AD

AD is primarily known as a neurodegenerative disease in which several brain phenotypes have been associated with its progression. Structural imaging evaluations with MRI consistently find that the entorhinal cortex is affected. Pathological characteristics include hippocampal volume loss or asymmetry, lower brain volumes and widespread loss of cortical thickness (Congdon & Sigurdsson., 2018). Additionally, abnormalities such as brain infarcts or white matter hyperintensities are standard features of AD, with some studies showing higher white matter hyperintensities in parietal and occipital lobes. Recent evidence has also suggested that the blood-brain barrier becomes compromised as part of the AD pathological cascade (Van De Haar et al., 2017).

Hippocampal atrophy is present in over 80% of AD cases (Hollands et al., 2016), but it is not clear why it is often the initial disease epicentre. In vivo, human brain imaging studies cannot achieve sufficient spatial resolution to examine cellular processes. Mice studies investigating mechanisms of pathology have found that when comparing strains of tau aggregate, the hippocampus was highly vulnerable to all strains (Mrdjen et al., 2019). Selective vulnerability can be explained by the cell types found in these hippocampal areas and may also be due to down-regulation of genes related to synaptic transmission and vesicular transport and up-regulation of inflammatory responses (Mrdjen et al., 2019). Mrdjen et al. (2019) found that specific neuronal subtypes were affected by tau in different disease stages, and the level of vulnerability also differed by stage. This work summarised cellular and molecular evidence for individual differences within neuroanatomical vulnerability for AD and supported the development of precise disease subtypes and phenotypes.

Additional neuroimaging studies have attempted to identify such subtypes within AD to implement precision medicine approaches more effectively. A systematic review by Ferreira et al. (2020) supported existing evidence for three clinically different subtypes of AD based on regional atrophy through post-mortem and neuroimaging data. These include typical AD, which affects the hippocampus

and association cortex, limbic-predominant AD, which primarily affects the hippocampus and hippocampal-sparing AD, which primarily affects the association cortex. The hippocampal-sparing subtype was found to have the earliest onset, with females and *APOE*-ε4 carriers more frequently having the limbic-predominant subtype. There were apparent differences in demographic variables between subtypes. Namely, hippocampal sparing was found more frequently in individuals with higher education levels and typical atrophy was found more frequently in individuals with lower levels of education, implicating lifestyle factors in differential brain MRI phenotypes in AD which tend to not be considered.

Evidence of heterogeneity has encouraged research questions on the influence of risk factors, neurogenetics and presenting phenotypes in AD development. It is thought that patterns of disease propagation are mediated not only by underlying neural architecture but many metabolic pathways (Filippi et al., 2020). The complexities of linking imaging to differences in neurological function are also difficult and have not been studied as extensively with consideration of subtypes; therefore, it is not yet understood whether these differences can predict cognitive decline or prognosis. It is unclear which brain MRI phenotypes and structures are significant to study in relation to early pathology and this has been an area of interest with inconsistent findings.

5.1.5 Cognitive phenotypes of AD

According to the National Institute of Aging in the US, individuals must meet specific criteria before a probable diagnosis of AD can be made. This includes cognitive impairment that cannot be attributed to a psychiatric disorder, which must be diagnosed through neuropsychological examination, and interferes with daily functioning affecting at least two of the following domains: the ability to acquire new information, impaired reasoning or poor judgement, impaired visuospatial abilities, language functioning or changes in personality or behaviour. The primary differentiation from MCI is that these affected domains must interfere with functioning in daily activities (McKhann et al., 2011). The distinction between MCI and amnestic MCI (aMCI) is often made in the literature as both structural differences and disease prognosis has been found to differ between the two groups consistently (Seppälä et al., 2010; Counts et al., 2016;

Lui et al., 2013). A longitudinal study by Wilson et al. (2011) characterised cognitive changes that may occur during prodromal phases of AD based on 19 neuropsychological tests and found that entorhinal cortex tends to be the first brain region affected in those who go on to develop AD. The study also found that MCI individuals who developed AD showed a sharp decline in cognitive abilities in the 5 to 6 years preceding diagnosis, in comparison to MCI individuals who did not progress to AD. They found that those presenting with amnestic MCI progressed around two years faster than non-amnestic MCI individuals, suggesting cognitive ability is an early marker for potential impairment and AD. Changes in white matter microstructure have also been implicated in AD where processing speed is affected (Salami et al., 2012.

5.1.6 Previous studies: genetic risk of AD and cognition or brain MRI

The first use of a PGR score for AD was published in 2015 by Escott-Price et al., which used GWAS results from the International Genomics of Alzheimer's Project to investigate the prediction accuracy of weighted genetic variants. They were able to predict conversion with a ROC curve of 78%, which included the *APOE* genotype. In recent years, studies have used PGR scores of AD for phenotype association to identify possible causal determinants, improve risk prediction and novel diagnostic approaches. One study by Xiao et al. (2017) investigated the influence of AD risk alleles on structural MRI measures, including hippocampal function and cognitive measures (N=231). This study found reduced brain function and metabolism in the hippocampus measured by PET and fMRI in healthy individuals with high PGR scores. They found no association of PGR score with cognition measured by a neuropsychological battery. However, Mormino et al, 2016 also studied healthy participants (N=1,322) and found elevated PGR was associated with worse memory ($p = 0.002$) and smaller hippocampus ($p = 0.002$) at baseline, as well as greater longitudinal cognitive decline (memory: $p = <0.001$, executive function: $p = 0.01$) and clinical progression to AD ($p < 0.001$).

Other studies have found that genetic risk for AD was associated with several structural brain volume metrics, suggesting additional brain imaging markers may also be clinically relevant. For example, Chauhan et al. (2015) looked at 24 novel risk variants (n=8,175 to 11,550). The study investigated the association

with MRI- markers of structural brain ageing in older, healthy participants and found that a weighted genetic risk score, novel AD genetic risk variants were associated with smaller brain volumes, particularly hippocampal volume.

However, one study by Harrison et al. (2016) found that both weighted and unweighted risk scores for AD did not associate with any differences for brain regions at baseline but was associated with hippocampal thinning two years after baseline in cognitively healthy older adults. This suggests PGR may be a valuable tool for predicting conversion to AD or pathological trajectories. However, the study considered *APOE* alleles within weighting, and it may be meaningful to look at genetic risk beyond *APOE*, particularly as many studies investigate the influence of *APOE* ε4 on risk factors and interactions with genetic factors. Louge et al. (2019) demonstrate a similar prognostic value of genetic risk on cognitive functioning. They found that higher PGR for AD (n=1,176) was associated with higher odds of having aMCI than being cognitively normal (OR: 1.36 - 1.43, $p<0.2$).

Numerous heritable factors play a role in risk and aetiology, and some studies suggest that genetic loci showing most effects within GWAS (i.e., at the top of the hit list) may mediate effects on phenotype outcomes. This brings to question whether the interactions of genetic risks may also be of interest. Marioni et al. (2017) stratified individuals by risk and found that individuals with high compared to low PGR scores for AD (top and bottom 5% of the distribution) show significantly lower cognitive functioning. However, when comparing cognitive differences by *APOE* status differences were larger; ε4ε4 alleles compared to ε3ε3 showed 1.2 fewer points for processing speed and 1 point fewer for memory, $p<.05$. Lupton et al. (2016) found that PGR (including the ε4 risk allele) was associated with reduced hippocampal and amygdala volume in MCI and healthy populations. However, *APOE* ε4 was associated with hippocampal and amygdala volume in AD and MCI but not in healthy older adults. Suggesting genetic risk including and excluding *APOE* ε4 status may both be valuable in identifying those at risk. A study by Lyall et al. (2019) found that although the ε4 genotype did not interact with lifestyle factors, there were suggestive interaction results that men were more vulnerable to the ε4 genotype in terms

of cognition. Studying such interactions with other genetic risks or MRI measures may be of interest.

5.2 Study rationale

Previous studies investigating associations of PGR of AD and MRI and cognitive functioning measures have focused on specific aspects of hippocampal function or structure in small sample sizes and with relatively poor 'controlling' for confounding variables like smoking. The sample size of this study and the discovery GWAS sample are bigger than previous studies, with the largest previous study using PGR of AD populations and MRI data to our knowledge being Marioni et al., (2017) n=3,495. The present study includes n=32,790 individuals for analyses. Genetic interaction between loci showing largest effects in GWAS have not been extensively studied, and it is known that *APOE* ε4 poses the most significant genetic risk of AD with potential gene-gene and gene-environment interactions of ε4 (Lyall et al., 2019; Lupton et al., 2016). Therefore, interaction of the *APOE* ε4 alleles and PGR will be analysed by including *APOE* status as an interaction term as either 0/1/2 depending on the number of ε4 alleles present.

Summary statistics from Kunkle et al., GWAS meta-analyses (case: n=30,344, control: n=52,427) will be used to create polygenic risk scores, and analyses will be carried out using 32,790 participants in the UK Biobank. This will be the largest study to date investigating PGR of AD in relation to brain MRI and cognitive functioning. Dependent variables include 12 structural MRI volumes and 6 cognitive measures. We hypothesise that high genetic loading based on a polygenic risk score of 21 loci implicating Aß, Tau, immunity and lipid processing in AD is associated with worse brain MRI and cognitive outcomes in healthy older adults within the UK Biobank cohort.

5.3 Methods

5.3.1 UK Biobank data

In 2014 at the first imaging visit, ~40K UK Biobank baseline participants were invited and recruited to return for the very first brain and body imaging and further follow up assessments. Participants used for the present analyses were filtered from this instance. At the time of analysis, there were 32, 790 individuals with MRI data who met inclusion criteria and were included for this chapter.

5.3.2 Generating polygenic risk score and genetic variants for AD

PGR scores were generated for all individuals with genotype data available in UK Biobank using LDpred. Summary statistics from a GWAS meta-analysis using 46 total datasets (cases n=30,344 and cognitively normal controls n=52,427) was used. 21 genome-wide-significant loci (6, 578,321 SNPs) associated with AD (onset age >65 years) were included to calculate weighted scores. The two SNPs, rs429358 and rs7412, used to define *APOE* status were removed from summary statistics and were not carried forward to calculate polygenic risk scores. APOε4 allele dose was included as a separate variable for each individual coded as 0, 1 or 2 according to the number of ε4 alleles and individual had. Once polygenic risk scores were calculated, scores close to 0 indicated lowest genetic loading and negative scores indicated greater genetic loading. This was standardised to Z-scores, i.e. mean=0, SD=1.

5.3.3 Genotyping

Some individuals were not included for genetic QC. Individuals not carried forward for genetic QC had non-white British ancestry, self-report vs. genetic sex mismatch, putative sex chromosomal aneuploidy, excess heterozygosity, and missingness rate > 0.1. Genetic QC included SNPs with a MAF greater than 0.01 and SNPs were tested for HWE in both validation and Biobank cohorts in which SNPs that had a p-value less than 0.001 were not included, this left 6, 578,321 SNPs included in the PGR score calculation which was calculated using LDpred. LD was calculated using 1000 unrelated Biobank participants who were not used in final analyses but passed genetic QC, this was to prune the SNP set used for

the PGR score for minimal LD. Further details about genotyping methods used in UK Biobank are outlined in chapter 2.

5.3.4 Brain MRI phenotypes

12 MRI volumes considered to have a priori evidence as major substrates of cognitive decline or AD pathology were included as outcome measures in this study: grey matter, white matter, white matter hyperintensity, whole brain, left hippocampus, right hippocampus. Additionally, segmented regions of each left and right hippocampus (head, tail and body) were examined. MRI variables were converted to Z-scores for interpretation and comparison.

5.3.5 Cognitive phenotypes

Cognitive measures used in analyses included: fluid intelligence (reasoning), prospective memory, numerical memory, reaction time, trail making (a+b), symbol digit substitution. Pairs matching was not included in this study as a paper by Lyall et al., (2020) assessing reliability and validity of UK Biobank cognitive tests demonstrated the reliability of this measure within the UK Biobank cohort showed little test-retest reliability (r=0.19). Further details can be found in Chapter 2: Methodology.

5.3.6 Exclusions

There were 39, 755 participants with *APOE* genotype and brain MRI data. We excluded participants with non-white British ancestry, self-report vs. genetic sex mismatch, putative sex chromosomal aneuploidy, excess heterozygosity, and missingness rate > 0.1. We also removed participants who reported a neurological condition at baseline. This left n = 32, 790 individuals to include for analyses.

5.3.7 Analyses

The predictor variables for these analyses include *APOE* ε4 allele dose and PGR of AD. Dependent variables include brain MRI and cognitive phenotypes. Regression models for all analyses were both partially adjusted and fully adjusted for each MRI and cognitive variable of interest. Partially adjusted

models controlled for age, sex, BMI, genotyping chip and 8 UK Biobank principal components which control for population stratification. Fully adjusted models were controlled additionally for Townsend deprivation score, education, smoking (as measured by pack per year as a proportion of lifetime exposure) and *APOE* ε4 dose. Interactions between PGR score and *APOE* ε4 dose were then analysed in for each dependent variable using an interaction term (PGR***APOE*) to assess if there were differences in PGR association with dependent variables for different *APOE* ε4 allele doses. All betas are standardised betas for ease of comparison, and bold results indicate significance at $p<0.05$.

5.4 Results

5.4.1 Demographic statistics

As detailed in the methodology, 32, 790 participants were included in analyses. PGR scores were calculated for all individuals and divided into tertiles for the purpose of table 5.1 showing demographic statistics by tertile. Tertile 1 represents individuals with the highest genetic loading.

Table 5.2 shows demographic information by *APOE* dose. The *APOE* dose split for all individuals included in analyses were: 73% had 0 'dose' of *APOE* ε4, 24% had 1 'dose' of the ε4 allele and 4% had two 'doses' of the ε4 allele. Individuals with two ε4 alleles showed the highest proportion of females, lowest BMI and lowest level of education.

Table 5.1 PGR tertiles and demographic statistics

PGR tertile	N	Sex % female	BMI (kg/m2)	Age	Education (% degree)	Smoking (pack/years)	Townsend	*APOE* ε4 dose 0 / 1 / 2
1	11,152	53%	26.47 (4.32)	63.85 (7.54)	26.20%	0.37 (0.31)	-2.00 (2.1)	73% / 24% / 3%
2	10,919	52%	26.48 (4.39)	63.76 (7.49)	26.32%	0.39 (0.3)	-1.99 (2.64)	73% / 23% / 4%
3	10,719	53%	26.5 (4.37)	63.86 (7.43)	25.22%	0.39 (0.3)	-2.03 (2.61)	72% / 24% / 4%

Table 5.2 Demographic statistics by *APOE* dose

APOE dose	N	Age	Sex (% female)	BMI (kg/m2)	Education (% degree)	Smoking (pack/years)	Townsend
0	23,830	63.77 (7.55)	52%	26.52 (4.36)	57%	0.386 (0.314)	-1.899 (2.714)
1	7,761	63.32 (7.41)	54%	26.38 (4.39)	55%	0.377 (0.308)	-1.885 (2.717)
2	1,199	63.04 (7.26)	55%	25.99 (4.09)	54%	0.351 (0.301)	-1.981 (2.633)

Table 5.3 Demographic statistics by male and female genetic sex

Male vs Female	N	Age	BMI (kg/m2)	Smoking (pack/years)	Townsend	Education (% degree)
Female	17,580	63.05 (7.38)	26.07 (4.71)	0.35 (0.28)	-1.86 (2.71)	57%
Male	15,210	64.34 (7.63)	27.02 (3.89)	0.41 (0.33)	-1.94 (2.71)	56%

5.4.2 PGR of AD associations

Table 5.4 shows significant deleterious associations with PGR were found for both left (β = -0.146, p = 0.002) and right (β = -0.128, p =0.026) hippocampal volumes when models were partially adjusted. Left hippocampal volume was significantly associated (β = -0.118, p =0.020) with PGR when additional covariates were added to the model, but not right hippocampal volume (β = -0.071, p =0.099), this was the only significant association with PGR when the model was fully adjusted. White matter hyperintensities showed significant interactive effects of PGR and *APOE* status when partially adjusted (β = 0.034, p =0.002) but not when fully adjusted (β = 0.062, p =0.972). There were no significant interactive effects for any structural MRI volumes in the fully adjusted model.

Table 5.5 shows PGR of AD associations with hippocampal subdivisions. Results show lower hippocampal subdivisions in partially adjusted models for left hippocampal head (β = -0.036, p =0.003), body (β = -0.056, p =0.002), right hippocampal tail (β = -0.071, p =0.167) and right hippocampal head (β = -0.046, p =0.023). When fully adjusted for covariates, all associations remained between PGR and left hippocampal head (β = -0.014, p =0.017), body (β = -0.069, p =0.002) and tail (β = -0.027, p =0.016), and also for right hippocampal head (β = -0.017, p =0.044). There were no significant interactive effects between PGR and *APOE* dose for hippocampal subdivision volumes in this model when partially or fully adjusted for covariates.

Table 5.6 shows 6 cognitive measures were significantly associated with genetic risk of AD when the model was partially adjusted for covariates. These included fluid intelligence (β = -0.066, p= 5×10^{-6}), prospective memory (β = -0.067, p =0.006), numerical memory (β = 0.232, p =0.043), reaction time (β = 0.022, p =0.003), trail making (β = 0.024, p =0.002) and symbol digit substitution (β = -0.114, p =0.001). One measure remained significantly associated with PGR when additionally adjusted for social deprivation, education, smoking and *APOE* dose. This was fluid intelligence (β = -0.080, p =0.002).

Table 5.4 Associations between PGR of AD and MRI volumetric measures

	Partially adjusted				Fully adjusted			
	Polygenic risk AD		PGR*_APOE_ ε4		Polygenic risk AD		PGR*_APOE_ ε4	
	β	p	β	p	β	p	β	p
Left hippocampal volume	**-0.146**	**0.002**	-0.127	0.620	**-0.118**	**0.020**	-0.121	0.447
Right hippocampal volume	**-0.128**	**0.026**	-0.111	0.729	-0.071	0.099	-0.185	0.663
Whole brain Volume	-0.055	0.201	-0.050	0.475	-0.049	0.075	-0.091	0.743
White matter volume	-0.058	0.351	-0.051	0.801	-0.022	0.115	-0.013	0.158
White matter hyperintensity volume	0.026	0.054	**0.041**	**0.006**	0.007	0.549	0.073	0.732
Grey matter volume	-0.173	0.393	-0.0562	0.933	-0.051	0.904	-0.081	0.472

Table 5.5 Associations between PGR of AD and hippocampal subdivisions

	Partially adjusted				Fully adjusted			
	Polygenic risk AD		PGR*_APOE_ ε4		Polygenic risk AD		PGR*_APOE_ ε4	
	β	p	β	p	β	p	β	p
Left Hippocampal head	**-0.036**	**0.003**	-0.034	0.39	**-0.014**	**0.017**	-0.041	0.079
Left Hippocampal body	**-0.056**	**0.002**	-0.087	0.143	**-0.069**	**0.002**	-0.043	0.300
Left Hippocampal tail	-0.071	0.167	-0.021	0.080	**-0.027**	**0.016**	0.022	0.124
Right Hippocampal head	**-0.046**	**0.023**	-0.033	0.080	**-0.017**	**0.044**	-0.012	0.572
Right Hippocampal body	-0.029	0.063	-0.025	0.816	-0.031	0.074	-0.01031	0.126
Right Hippocampal tail	-0.043	0.935	-0.031	0.815	-0.095	0.932	-0.0230	0.113

Table 5.6 Associations between PGR of AD and cognitive function

	Partially adjusted				Fully adjusted			
	Polygenic risk AD		PGR*_APOE_ ε4		Polygenic risk AD		PGR*_APOE_ ε4	
	β	p	β	p	β	p	β	p
Fluid intelligence	**-0.066**	**5×10^{-6}**	-0.112	0.704	**-0.080**	**0.002**	-0.180	0.772
Prospective memory	**-0.067**	**0.006**	-0.039	0.458	-0.369	0.460	-0.511	0.195
Numerical memory	**-0.232**	**0.043**	-0.023	0.986	-0.039	0.178	-0.056	0.954
Reaction time	**0.022**	**0.003**	0.058	0.811	0.014	0.226	0.052	0.201
Trail making	**0.073**	**2×10^{-6}**	-0.008	0.0689	0.026	0.112	0.006	0.430
Symbol digit	**-0.114**	**0.001**	-0.012	0.884	-0.038	0.090	-0.021	0.889

5.4.3 *APOE* ε4 dose associations

Table 5.7 shows *APOE* significant associations for both MRI and cognitive measures. ß shows standardised betas reflecting per *APOE* ε4 allele increase. Results showed significant associations between *APOE* dose and cognitive functioning for four measures when partially adjusted, but none remained significantly associated when fully adjusted. These include fluid intelligence (ß = -0.064, p =0.012), prospective memory (ß = -0.071, p =0.032), trail making (ß = -0.034, p =0.024) and symbol digit substitution (ß = -0.173, p =0.032). 6 MRI measures showed significant association with *APOE* dose when fully adjusted for covariates. Left hippocampal volume (ß = -0.069, p =0.024), right hippocampal volume (ß = -0.079, p =0.002), white matter hyperintensity volume (ß = 0.063, p =0.006), left hippocampal body (ß = -0.094, p =0.048) right hippocampal body (ß = -0.067, p =0.046) and right hippocampal tail (ß = -0.076, p =0.042).

Table 5.7 Associations between *APOE* ε4 dose and MRI and cognitive measures

	APOƐ4 dose partially adjusted		APOƐ4 dose fully adjusted	
	β	p	β	p
Whole brain volume	-0.223	0.301	-0.094	0.138
White matter	-0.240	0.347	-0.074	0.206
Grey matter	-0.089	0.430	-0.022	0.302
Left hippocampal volume	**-0.064**	**0.019**	**-0.069**	**0.024**
Right hippocampal volume	**-0.039**	**0.002**	**-0.079**	**0.002**
White matter hyperintensity volume	**0.042**	**0.002**	**0.063**	**0.006**
Left hippocampal body	-0.026	0.065	**-0.094**	**0.048**
Right hippocampal body	-0.015	0.347	**-0.067**	**0.046**
Right hippocampal tail	-0.026	0.117	**-0.076**	**0.042**
Fluid intelligence	**-0.064**	**0.012**	-0.069	0.223
Prospective memory	**-0.071**	**0.032**	-0.086	0.208
Numeric memory	-0.053	0.085	-0.025	0.648
Reaction time	0.037	0.063	0.042	0.700
Trail making (a+b)	**0.034**	**0.024**	0.059	0.056
Symbol digit	**-0.173**	**0.032**	-0.244	0.155

Note. β = standardised betas reflecting per *APOE* ε4 allele increase. Bold = uncorrected significant p<0.05. Models are partially adjusted for age, BMI, sex, genotyping chip, 8 genetic principal components and fully adjusted for additional smoking, education and social deprivation.

5.5 Discussion

The current study examined the genetic risk of AD in 32, 790 healthy adults within the UK Biobank and found that PGR of AD was significantly associated with MRI volumes of the whole left hippocampus, left hippocampal head, body and tail, right hippocampal head and fluid intelligence when fully adjusted for age at the time of assessment, genotyping chip and batch, 8 UK Biobank principal components, sex, BMI, smoking, education, social deprivation and *APOE*4 dose. These results suggest genetic risk variants for AD may be able to indicate the earliest signs of pathology prior to cognitive problems and that there is potential for PGR based AD risk assessment to be made before the onset of symptoms. There was no evidence in these analyses to suggest the association of PGR with MRI volumes or cognition is dependent on *APOE*4 dose in healthy adults.

5.5.1 Interpretation

Many AD studies report baseline presence of WMH on MRI as a correlate of cognitive decline or future MCI status (Jacobs et al., 2012; Tosto et al., 2015) in healthy adults, whereas these analyses did not find evidence that genetic risk of AD was associated with WMH. A study using the Framingham Offspring cohort by Bangen et al. (2018) found that hippocampal volumes, but not WMH volume, were associated with the conversion of healthy adults to MCI. Although the present analyses did not use longitudinal data, there was similarly a lack of evidence for the presence of WMH in healthy adults who who may have an elevated genetic risk for AD or MCI. The similarities between these findings and our results indicate that the earliest signs of pathology in healthy adults may not consistently implicate WMH. One explanation for the commonly found presence of WMH could be vascular risk factors and vascular diseases that tend to be associated with the development of WMH and comorbid with dementias (Attems & Jellinger, 2014). A study by Armstrong et al. (2020) found that candidate loci for WMH were implicated in stroke, vascular and neuronal functions, but not dementia when conducting a GWAS for periventricular WMH and deep WMH. This raises the concern of differentiating AD markers from non-AD markers such as vascular pathology that may still contribute to AD. However, this is difficult with

observational studies as AD refers to an aggregate of neuropathological changes assessed postmortem.

One way this may be better understood is by integrating additional types of imaging data such as connectomics, which can give more specific pathological information about why some affected regions show further neurodegeneration changes and others do not. One theory put forward in a paper by Filppi et al. (2020) suggested that once a disease-epicentre is established, further propagation of pathology or integrity changes may depend on whether the integrity of white matter pathways has been compromised. Therefore disease pathology can spread along discrete brain networks. In the context of the present study, this may explain why we found disease epicentres such as the hippocampus showing differences with higher genetic risk where neurodegeneration of structural connectivity may not have yet begun to occur. In these analyses, individuals at high genetic risk may represent initial stages of AD, with risk scores reflecting primarily affected phenotypes. Mainly as the UK Biobank sample is generally healthier than the general population with a younger sample than an older adult population, we may expect only to find early signs of pathology.

Further neurodegeneration or cognitive decline from initial pathology may depend on an individual's cerebrovascular health, e.g., Lyall et al. (2017) found that an increasing number of cardiometabolic comorbidity was associated with worse cognitive abilities. However, no discrete biological pathways or events reliably lead to AD, and variants identified by PGR scores are limited by the current diagnostic criteria and specificity of AD. Studies investigating clinically typical and atypical subtypes of AD suggest specific sub-populations present differing pathologies and genetic risk variants, e.g. SNAP populations presenting with lower tau levels may not be identified using a risk score that includes tau related variants (Dani et al., 2017). A study by Mohanty et al. (2020) proposed harmonizing neuroimaging and subtyping methods for AD. There is further evidence that subtypes of AD have differing genetics. Some studies have identified that *APOE* ε4 status can differ amongst subtypes; Leonenko et al. (2019) found that *APOE* ε4 contribution in polygenic risk predictions may indicate amyloid deposition. Other risk loci for the PGR included variants associated with

amyloid precursor, cholesterol transport, calcium homeostasis, tau binding and immune response.

5.5.2 PGR and *APOE* interactions

No PGR × *APOE* interactions were found in fully adjusted models in these analyses. Left hippocampus and left hippocampal body volumes showed significance for both *APOE* dose and PGR separately in fully adjusted models, but no evidence of interaction was found for these metrics by the model. However, genetic interactions for complex traits are not well understood and there may have been potential methodological limitations as to why no interactions were found in this study. It is also important to remember when considering associations and interactions that the association of *APOE* with phenotypes may not necessarily indicate AD, numerous studies have shown that *APOE* is a risk factor for cerebral angiopathy, Lewy body dementia, cerebrovascular diseases, and multiple sclerosis (Verghese et al., 2011).

5.5.3 *APOE* associations

APOE dose was significantly associated with 6 MRI measures: left hippocampal volume, right hippocampal volume, WMH, left hippocampal body and right hippocampal body and tail. These results are consistent with current literature that ε4 carriers appear to be at increased risk of hippocampal atrophy and worse WM microstructural integrity (Bangen et al., 2018). However, no cognitive measures showed *APOE* dose associations in fully adjusted models. It is possible that may in part be explained by the fact that anatomical substrates of cognitive decline will occur a number of years before cognitive symptoms show, even when brain ageing is accelerated by the *APOE* risk factor (Hedden and Gabrieli, 2005; Vernooij and Smits, 2012). *APOE*

5.5.4 Limitations

The genetic associations observed in this study may be limited by the methodology of calculating genetic risk scores and modelling assumptions. There is no unified approach to efficiently calculating PGR scores, with variations of methods differently accounting for linkage disequilibrium, beta shrinkage, and GWAS p-value thresholding. Residual signal may come from the *APOE* genotype

within the AD-PGR scores we calculated, as we did not remove all SNPs in LD with the *APOE* ε4-defining SNPs (rs429358 and rs7412) although we did subsequently control for this genotype and therefore, to some extent SNPs in LD. In addition, we could not replicate genetic effects in an independent cohort. This is due to challenges finding an appropriately phenotyped cohort not included within the original GWAS meta-analysis; replication of gene and structural imaging associations is a scientific priority going forward. However, recently published recommendations by Wand et al. (2021) were considered, in which reporting standards have been met.

Another limitation of this study is with cognitive testing in which the prospective memory measure was associated with PGR and *APOE* interactions only when partially adjusted. UK Biobank's prospective memory measure showed a ceiling effect as it was measured as a binary variable, i.e., completed or not completed, with 98% of individuals having completed the task.

A major limitation of these findings is due to demographics of the UK Biobank cohort, who are overall less likely to have health conditions, are more educated and live in less socioeconomically deprived areas. Further selection bias may be found for MRI data as data collection is time-consuming. Additionally, the findings of this study cannot be generalised to populations outside of White European ancestry. Genetic risk variants contributing to AD are known to differ among populations. A study by Pham et al., (2018) looking at primary care data in the UK, reported that African ethnic groups have a higher incidence of AD than European individuals; however, they are underdiagnosed in comparison. A review by Abondio et al. (2019) showed that the distribution of *APOE* allele frequencies varies across the globe, illustrating the variance of genetic factors that require consideration if findings are intended to benefit those who show to be most at risk. Cohort studies of different ethnic groups are also necessary to fully understand AD's prevalence and biological pathways.

5.5.5 Future research

The current analyses use cross-sectional data; future analyses could use longitudinal data to assess the validity of a PGR score to identify whether genetic risk showed early cognitive and brain MRI markers in those who went on

to develop AD. It may also be interesting in longitudinal analyses to examine the role of genetics or *APOE* ε4 in the progression rate for those who develop AD. The data is currently conflicting regarding whether ε4 vs ε3 influences the progression rate. It has been hypothesised that the ε4 genotype may be associated with faster decline due to the association of ε4 with additional diseases, creating additive or interactive pathologies. It is also well understood that family history is an informative marker of genetic risk in AD, often considered in a clinical context. It may be useful to test associations of PGR and AD by proxy with PGR. Marioni et al. (2018) showed that self-reported parental AD was a valid proxy for an AD genetic study.

Studies of PGR and AD in both healthy individuals and individuals who convert demonstrate PGR as a tool with the potential to identify individuals at risk before presenting clinical symptoms. Determination of such susceptibility can be beneficial for conditions where preventative measures have the potential to delay onset. The National Institute on Aging and Alzheimer's Association Research Framework has recommended a shift toward a biological definition, and the use of biomarkers for in vivo Alzheimer's diagnosis to identify brain markers or profiling cognitive deficits might help gain a better picture of those at risk. This relationship between genetic risk and biomarkers of AD can provide deeper insights into disease pathology and overall risk.

5.5.6 Conclusion

Chapter 5 provided evidence to suggest PGR of LOAD is associated with brain structure in healthy individuals with a mean age of 63, suggesting PGR may be a useful tool for those in need of early identification for intervention treatments.

Chapter 6: Associations between PGR of AD and blood biomarkers

6.1 Introduction

Genetic markers, neuropsychological examinations, and brain imaging approaches to detect and diagnose high-risk individuals of AD, although extremely useful, are expensive and time-consuming, whereas examination of blood and fluid biomarkers are clinically accessible. CSF measures of ß-amyloid and tau are the most common clinical methods used to diagnose probable AD with high specificity and sensitivity (Khoury & Ghossoub, 2019). However, collection of CSF is challenging to obtain as it is invasive due to a lumbar puncture and expensive to carry out compared to blood collection, which is much more widespread in clinical practice. Although both Aß and tau have been studied for decades in both CSF and blood, they do not occur in isolation or necessarily indicate AD. For example, tau levels are high where other types of neuronal injury have occurred, such as hypoxia, and research is conflicting regarding how closely blood Aß levels reflect Alzheimer's disease in the brain (Toledo et al., 2013). It is necessary to utilise additional biomarkers to elucidate the heterogeneity of specific disease pathways and trajectories of AD, moving towards a precision medicine approach within AD. Investigating blood biomarkers has been of particular interest as common biomarkers are taken in primary care settings and with the ability to reflect plasma levels of proteins, peptides, and lipid contributions from different tissues (Zetterberg., 2019).

6.1.1 Blood biomarkers

Blood biomarkers have the potential to improve detection and reduce costs in primary care settings and within the earliest stages of disease. For this reason, identifying novel fluid biomarkers in CSF and blood has been the focus of AD research in recent years, in which several emerging biomarkers have been proposed, although not yet validated. One such example includes increased CSF concentrations of neurofilament light (NFL), a biomarker identified via immunoassays that represent axonal damage in neurological disorders (O'Bryant, 2017). This biomarker is now being considered as a prognostic AD biomarker.

A recent landscape analysis (Hampel et al., 2018) found that out of 1,404 studies of blood-based biomarkers for AD, 34% were on markers of relevant mechanisms such as inflammation, mitochondrial dysfunction or microvascular injury, 18%

were on Aß and tau, 29% studied biomarker panels, and 19% were looking at genetic markers such as *APOE*. Several commonly studied biomarkers have been implicated in AD. For example, studies looking at protein levels in AD have found neurogranin, phosphodiesterase, albumin, cholesterol, LpA beta-secretase and rheumatoid factor to be slightly to moderately elevated in AD compared to controls (Wojsiat et al., 2017; Huynh & Moham, 2017; Proitsi et al., 2017; Blennow et al., 2010; Sanfilippo et al., 2016). It is hypothesised that lipids and lipoproteins also play a more significant role in developing AD than previously considered due to lipid involvement in the blood-brain barrier (BBB), amyloid precursor processing, myelination, inflammation responses and oxidation (Chew et al., 2020). Both AIBL and ADNI1 flagship studies have been involved in lipidomic analyses where lipid profiles of AD patients were used in longitudinal analyses to validate lipid models, which improved disease classification and prediction (Huynh et al., 2020). Previous lipidomic studies have suggested dysregulation of lipid profile early in the development of AD and associated cognitive impairment. For example, Kunkle et al. (2019) found that several lipid metabolism pathways were implicated, consistent with other studies (Reitz, 2013).

Many studies have investigated blood-based metabolic changes using a case-control design with the aim of stratifying patient populations by pathophysiology. However, studying blood biomarkers in a genetic context has not been as extensively examined. Additionally, previous studies and cohorts included in research looking at blood biomarkers in the context of AD include patient populations or memory clinics. However, when seeking to validate markers within a general population, it is necessary to carry out exploratory and validation analyses on wider populations. For this reason, the UK Biobank is a valuable cohort for exploratory analysis.

6.1.2 Sex differences in AD and biomarkers

Age, sex and ethnicity are all variables that have been shown to influence the association of *APOE* e4 genotype with Alzheimer's and are also essential to consider when carrying out exploratory analysis for blood biomarkers (Farrer et al., 1997; Zhao et al., 2020). Biomarker profiles are understood to differ between men and women in the general population, in which sex differences are

most apparent for biomarkers of adiposity, cardiovascular stress and inflammation (Lew et al., 2017). When looking at sex differences in AD, Mosconi et al. (2017) found that the preclinical phase of AD is earlier in the lifetime for women than it is for men. One study by Lau et al. (2019) looked at 71 circulating blood biomarkers between men and women, for which they found that 61 biomarkers differed significantly, including adipokines, inflammatory markers and c reactive protein (CRP). When looking at associations with CVD, there was a sex interaction for ApoB. Research investigating sex differences within the progression of AD indicates sex-specific patterns of disease manifestation, implicating estrogen-related systems (e.g. menopause), sex-genotype interactions (e.g. *APOE*), cardiovascular-specific risk factors and gender-specific risk factors (Ferretti et al., 2018).

6.1.3 Study rationale

Among genetic studies of AD, there is some research investigating the role of genetic risk via *APOE* e4 genotype in the context of biomarkers. One study by Ferguson et al. (2020) looked at the *APOE* e4 genotype in the UK Biobank in relation to circulating blood biomarkers and found a range of associations between genotype and blood biomarkers in non-demented participants. In chapter 4 of this thesis, we have previously found that our PGR score of AD may indicate AD-related pathologies outside of the brain and *APOE* status. This study will conduct a systematic, hypothesis-free analysis of non-*APOE* PGR of AD for blood biomarkers in the UK Biobank in healthy adults.

6.2 Methods

6.2.1 Biomarker collection and processing

Biomarker levels were analysed in UK Biobank from serum and red blood cell samples at baseline from UK Biobank participants between 2006 – 2010 which included biomarker samples. This study investigated baseline measure of 30 blood biomarkers for 502, 536 individuals. This document outlines quality control approaches that were undertaken to mitigate random errors and bias in

biomarker samples: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf.

30 blood biomarkers of interest were used as dependent variables in these analyses including: low density lipoprotein (LDL), high density lipoprotein (HDL), lipoprotein a (LpA), oestradiol, phosphate, rheumatoid factor, sex-hormone-binding-globulin, total-bilirubin, direct-bilirubin, testosterone, protein, urate, c-reactive protein (CRP), triclycerides, gamma glutamyltransferase, dystatin C (*CysC*), vitamin D, ApoA, ApoB, creatinine , total cholesterol, calcium, urea, aspartate transaminase (AST), alanine aminotransferase (ALT), glucose, IGF-1, alkaline phosphate, albumin. Processing of very low levels of oestradiol and rheumatoid factor (RF) were recorded as missing, these values were recalculated as square root of the minimum stated detectable value if individuals had data for at least one other biomarker. Missing values for oestradiol and rheumatoid factor were only replaced if individuals had data for at least one other biomarker and were not coded as 'no data returned' or having unrecoverable aliquot problems. These 'missing' values were intended to be recoded to avoid them being zero and therefore not included, instead we recoded them conservatively to reflect very low levels. For this reason we chose to recode these values as the square root of the minimum stated detectable value. A square root transformation was applied instead of other methods because we wanted to record a low estimate.

6.2.2 Genetic risk score for AD

Non-*APOE* polygenic scores were calculated using LDPred and based on summary statistics from Kunkle et al. (2019) *APOE* PRS were generated for all individuals with genotype data available in UK Biobank using LDpred. Scores close to 0 indicated lowest genetic loading and negative scores indicated greater genetic loading. This was standardised to mean=0, SD=1. More details about how risk scores were calculated can be found in chapter 4.

6.2.3 Covariates and exclusions

Exclusions were made for those who reported a neurological condition at baseline and individuals who were not of white European ancestry. Covariates

included in these analyses were age, sex, Townsend deprivation, smoking (never vs ever), BMI, education, *APOE* dose (0/1/2) and CVD medication which was used as a proxy for CVD conditions. Medication use for hypertension, insulin and statins were included. This data was self-reported and taken at the first imaging visit alongside the imaging, cognitive and majority of covariate data (with the exception of Townsend deprivation score which was taken from baseline data). This was data-field 6177. Further information can be found here: https://biobank.ndph.ox.ac.uk/showcase/field.cgi?id=6177.

6.2.4 Statistical analyses

Minimally adjusted models were run controlling for: age, sex, assessment centre, PCs 1-8. Partially adjusted models were run controlling additionally for demographic factors including Townsend deprivation score, education and smoking. Finally, fully adjusted models controlled further for: BMI, CVD medication and *APOE* dose. Covariates for fully adjusted models differ from previous chapters as they were intended to account for covariates most relevant for biomarkers. A fourth model was run for all biomarkers using fully adjusted covariates with PGR*sex as an interaction term to study whether each biomarker showed an interaction with genetic risk.

6.3 Results

6.3.1 Sensitivity analyses

Biomarkers which were not normally distributed (testosterone, rheumatoid factor, oestradiol, gamma-glutamyltransferase, and c reactive protein) were log transformed and reanalysed: the resulting associations remained insignificant and we have reported the original estimates.

6.3.2 Demographic statistics

Table 6.1 and 6.2 show demographic statistics of individuals included in analyses, with table 6.2 showing demographics by sex. Table 6.3 shows general statistics used for each biomarker including number of individuals with available data for each.

Table 6.1 Demographic statistics

N	Age (years)	BMI (kg/m2)	CVD medication (% taking)	Smoking (% ever)	Townsend	*APOE*4 dose 0 / 1 / 2	AD PGR	AD PGR range
502 536	64.19 (7.75)	26.59 (4.46)	43 173 (8.59%)	3.7%	-1.29 (3.09)	72% / 25% / 3%	-1.25 (0.11)	-1.68 - -0.19

502, 536 individuals were included in these analyses for which covariates included age, BMI, CVD medication, smoking, education, Townsend deprivation and *APOE* dose. Age, BMI, Townsend and AD PGR is mean (SD).

Table 6.2 Demographic statistics by sex.

Male vs Female	N	Age (years)	BMI (kg/m2)	Smoking (never vs ever)	*APOE*4 dose 0/1/2	AD PGR	AD PGR range	Townsend
Female	273 399	63.50 (7.59)	26.15 (4.83)	3.13%	73% / 24% / 3%	-1.25 (0.11)	-1.66 - -0.19	-1.33 (3.04)
Male	229 137	64.93(7.84)	27.05 (3.98)	4.28%	73% / 24% / 2%	-1.26 (0.10)	-1.68 - -0.21	-1.25 (3.16)

Age, BMI, AD-PGR and Townsend are mean(SD). Males showed higher levels of smoking and AD-PGR range. Females showed a lower mean Townsend deprivation score.

Table 6.3 Biomarker descriptive statistics

Biomarker (nmol/L)	Mean	SD	Median	N
LDL	3.56	0.87	3.52	412 889
HDL	1.45	0.38	1.40	378 607
Triglycerides	1.58	0.97	1.39	413 156
HemoglobinA1C	36.13	6.77	35.20	410 807
LpA	44.65	49.21	21.10	330 832
Oestradiol	87.75	250.02	8.54	385 971
Phosphate	1.16	0.16	1.16	378 057
Rheumatoid factor	5.05	8.47	3.16	413 633
Sex-hormone-binding-globulin	51.63	27.78	45.27	375 079
IGF-1	21.36	5.47	22.04	412 684
Total bilirubin	9.13	4.42	8.07	411 877
Direct bilirubin	1.83	0.85	1.61	351 075
Testosterone	6.56	6.05	3.94	374 611
Protein	72.51	4.11	72.31	378 363
Urate	309.2	80.43	303.0	413 172
Vitamin D	48.61	21.11	46.80	394 928
CRP	2.60	4.36	1.33	412 764

Gamma-glutamyltransferase	37.39	42.08	26.30	413 447
Cystatin C	0.91	0.17	0.89	413 630
ApoB	1.03	0.24	1.02	411 547
ApoA	1.54	0.27	1.51	376 504
Creatinine	72.31	18.55	70.40	413 454
Total cholesterol	5.69	1.14	5.65	413 662
Calcium	2.38	0.09	2.38	378 673
Glucose	4.85	1.21	4.39	411 306
Urea	5.40	1.40	5.26	413 378
AST	26.23	10.66	24.40	412 093
ALT	23.07	11.65	22.47	375 618
Alkaline phosphate	83.67	26.45	80.40	413 675
Albumin	45.21	2.63	45.20	378 794

6.3.3 PGR of AD and biomarker associations

Table 6.4 shows PGR and biomarker associations. Model 1 was adjusted for age, sex, assessment centre, PCs. Model 2 was additionally adjusted for smoking, education, and Townsend deprivation score. Models 3 and 4 were fully adjusted for BMI, CVD medication and *APOE* dose. *CysC* was the only biomarker associated with PGR of AD in all three regression models. There was, however, no evidence for an interactive effect of sex and PGR of AD on the biomarker *CysC* in the fourth model. ApoB was associated with PGR when minimally (β =2.81, p=0.02) and partially adjusted (β =2.72, p=0.02) but not fully adjusted. Similarly to LDL in minimal (β= 8.95, p= 0.04) and fully adjusted models (β= 8.63, p= 0.05). Phosphate was associated with PGR in minimally adjusted model 1 (β =-1.31, p=0.09). 13 out of 30 biomarkers showed increased betas as each model controlled for further covariates, however there were no meaningful trends for change in p-values alongside increased beta value in any direction. These biomarkers include albumin, CRP, *CysC*, direct bilirubin, gamma glutamyltransferase, oestradiol, phosphate, rheumatoid factor, sex hormone binding globulin, total bilirubin, urate, urea, vitamin D. Additionally, in fully adjusted models, ApoA, HDL and LpA showed lower beta values than from the first two respective models in which significant association with CVD medication were also found. The largest effect sizes were seen with LDL (6.17- 8.95), HDL (6.17 - 8.95) urea (-7.36 - -8.21), testosterone (-2.80 - -5.79) and ApoA (0.01 - 5.95) none of which however showed significant associations with PGR of AD.

CRP (β= -0.82, p=0.02) creatinine (β= 3.21, p= 0.005), LpA (β=10.05, p= 0.08) and oestradiol (β= - 6.39, p= 0.05) showed a significant interactive effect between AD-PGR and sex. 12 out of 30 biomarkers showed an opposite (positive to negative or vice versa) association with PGR when testing for interactive effects with sex. The remaining biomarkers remained consistent in positive or negative directionality in their association with PGR of AD and interactive sex effects. This is due to 0 representing females and 1 representing males, meaning variables that were negatively associated with PGR*sex were higher for females than for males if the variable was positive in models 1 and 2 for the same biomarker.

Table 6.4 Associations between blood biomarkers and PGR of AD

	Minimally adjusted (model 1)		Middle adjusted (model 2)		Fully adjusted (models 3 & 4)			
	Polygenic risk AD		Polygenic risk AD		Polygenic risk AD		PGR * sex	
	β	p	β	p	β	p	β	p
Albumin	1.64	0.21	1.61	0.21	1.86	0.16	1.75	0.51
Alkaline phosphate	0.61	0.59	0.58	0.61	0.47	0.68	-2.19	0.34
ALT	0.64	0.55	0.76	0.28	0.62	0.41	0.18	0.49
ApoA	5.95	0.64	5.84	0.64	0.01	0.26	0.01	0.83
ApoB	**2.81**	**0.02**	**2.72**	**0.02**	1.99	0.11	-2.09	0.41
AST	0.56	0.27	0.57	0.26	0.62	0.25	-0.74	0.49
Calcium	4.89	0.29	4.79	0.31	4.79	0.33	3.87	0.69
CRP	0.11	0.55	0.11	0.54	-0.01	0.96	**-0.82**	**0.02**
Creatinine	-1.89	0.73	-0.21	0.70	-1.62	0.78	**3.21**	**0.005**

Cystatin C	**-1.26**	**0.03**	**-1.26**	**0.03**	**-1.47**	**0.013**	1.63	0.17
Direct bilirubin	1.34	0.69	1.35	0.70	1.86	0.16	0.17	0.52
Gamma glutamyltransferase	1.26	0.43	1.18	0.46	3.65	0.82	-1.78	0.59
Glucose	1.03	0.34	1.21	0.72	1.43	0.64	-0.78	0.39
HDL	5.72	0.74	5.40	0.76	1.76	0.31	3.59	0.92
IGF-1	-0.71	0.38	-0.28	0.46	-0.89	0.53	-0.12	0.49
LDL	**8.95**	**0.04**	**8.63**	**0.05**	6.17	0.18	-5.07	0.58
LpA	1.36	0.62	1.37	0.61	0.71	0.80	**10.05**	**0.08**
Oestradiol	-2.47	0.11	-2.36	0.12	-2.49	0.13	**-6.39**	**0.05**
Phosphate	**-1.31**	**0.09**	-0.01	0.10	-1.19	0.15	0.001	0.98
Protein	2.44	0.23	2.35	2.45	2.06	0.33	3.68	0.94

Rheumatoid factor	-0.82	0.06	-0.82	0.06	-0.74	0.11	-1.34	0.14
Sex hormone binding globulin	1.34	0.69	1.35	0.69	1.86	0.16	0.71	0.78
Testosterone	-3.09	0.83	-2.80	0.84	-5.79	0.69	-2.04	0.49
Total cholesterol	1.09	0.06	1.05	0.06	9.14	0.12	-5.51	0.64
Total bilirubin	1.34	0.69	1.35	0.70	1.86	0.16	-3.56	0.44
Triglycerides	0.78	0.23	0.97	0.31	0.58	0.48	0.94	0.73
Urate	2.17	0.48	2.10	0.50	3.78	0.99	-3.12	0.61
Urea	-7.36	0.21	-7.58	0.19	-8.21	0.18	-7.23	0.55
Vitamin D	-1.36	0.19	-1.34	0.19	-1.72	0.11	-1.28	0.55

6.3.4 Biomarker associations by sex

Table 6.5 shows five biomarker associations by female sex, and Table 6.6 shows the same biomarkers by male sex. This subset of biomarkers were shown as they were the only biomarkers that differed significantly between men and women, or showed evidence of sex interactions in previous models. When looking at CRP, creatinine, LpA and oestradiol by sex, creatinine, *CysC* and oestradiol were significantly negatively associated with higher genetic risk scores for females but not in men. *CysC* did not show an interaction by sex in the main models (partially or fully adjusted). When stratifying by sex, there was a stronger, significant association in females than in males who showed a weaker association in the same direction. Beta values were, however, still small.

When looking at CRP, creatinine, and oestradiol by sex, creatinine, *CysC* and oestradiol were significantly negatively associated with higher genetic risk scores for females but not for males.

Table 6.5 AD-PGR associations by female sex

Female	Minimally adjusted (model 1)		Middle adjusted (model 2)		Fully adjusted (model 3)	
	Polygenic risk AD		Polygenic risk AD		Polygenic risk AD	
	β	p	β	p	β	p
CRP	0.36	0.15	0.36	0.16	0.25	0.34
Creatinine	**-1.87**	**0.006**	**-1.88**	**0.006**	**-1.72**	**0.01**
Lipoprotein A (LpA)	-3.68	0.33	-3.71	0.33	-4.41	0.26
Cystatin C	**-0.02**	**0.01**	**-0.02**	**0.01**	**-0.02**	**0.004**
Oestradiol	**-5.59**	**0.04**	**-5.43**	**0.05**	**-5.68**	**0.05**

Table 6.6 AD-PGR associations by male sex

Male	Minimally adjusted (model 1)		Middle adjusted (model 2)		Fully adjusted (model 3)	
	Polygenic risk AD		Polygenic risk AD		Polygenic risk AD	
	β	p	β	p	β	p
CRP	-0.22	0.39	-0.21	0.39	-0.32	0.21
Creatinine	1.58	0.07	1.52	0.08	1.49	0.09
Lipoprotein A (LpA)	6.64	0.09	6.67	0.08	5.83	0.15
Cystatin C	-0.006	0.49	-0.006	0.50	-6.61	0.46
Oestradiol	1.01	0.07	1.00	0.08	1.42	0.10

6.4 Discussion

The current study investigated associations between circulating blood biomarkers and PGR of AD in 502, 536 healthy individuals. This study also investigated interactions with sex as there is not sufficient research exploring sex differences in relation to blood biomarkers for AD. In terms of genetic risk association with circulating biomarkers, decreased levels of *CysC* were associated with an increased polygenic risk of AD, particularly in women. This was the only biomarker associated with AD-PGR in all three models. There was also evidence to suggest lower levels of oestradiol and creatinine were associated with higher PGR in females but not in males.

6.4.1 Biomarker and AD-PGR associations

CysC was negatively associated with PGR in all three models (i.e., as genetic risk for AD increased, blood *CysC* tended to decrease). When split into male and female, this association only remained significant in women. Genetic and biochemical studies have supported the role of *CysC* expression in neurons and microglia, particularly pyramidal neurons that are also vulnerable to neurodegeneration in AD (Kaur & Levy, 2012). However, the genetic risk score in the present study did not contain the *CysC* gene, and the role of *CysC* in the brain is currently conflicting. When turning to biomarker studies, it is difficult to interpret the influence of *CysC* on the brain from existing research as it is primarily known as a renal biomarker which is typically high when kidney function is inadequate. However, it has been suggested that chronic kidney disease is itself a risk factor for cerebrovascular health, suggesting a potential mediator. This also suggests that lower levels of *CysC*, which are favourable for kidney function, may also be favourable for cerebrovascular health and cognitive functioning (Ling Lau et al., 2020). In line with this, higher levels of *CysC* are thought to contribute to increased neuronal vulnerability as CycC is expressed by neurons, astrocytes, and microglial cells, where increased *CysC* involvement has been implicated in neuronal death in AD (Deng et al., 2001).

By contrast, the results found in this thesis, particularly for women, showed that lower levels of *CysC* were associated with a higher genetic risk of AD. However, there are similarities with our findings and other studies of *CysC* in the context of AD, which show it plays a protective role in Alzheimer's by inhibiting aggregation of Aß (Matthews & Levy, 2016; Gauthier et al., 2011; Zhong et al., 2013). One reason for the differences in our findings by sex may be due to a complex interplay between sex-based differences, the vascular system and brain ageing. This is supported with results reported by Birgitta Werner et al. (2014), who found that male sex and vascular risk factors influenced levels of *CysC* in those without diabetes or vascular disease. Although it has been suggested that the association between *CysC* and age is non-linear, Birgitta Werner et al. (2014) found sex-specific effects of ageing for *CysC*. Taken together, these findings suggest there may be a nuanced role of *CysC* in the ageing brain that is difficult to explain with the results in this analysis.

When looking at biomarker and PGR associations, variability of the beta values for fully adjusted models compared to respective beta values for minimally and partially adjusted models was notable. This brings to question the role BMI, CVD or *APOE* may play in this population in modifying the association between polygenic risk of AD and blood biomarker levels. This was seen particularly with ApoA, gamma-glutamyltransferase, HDL, LpA, testosterone, total cholesterol and urate. One avenue by which these covariates may influence the association between genetic risk of AD and these circulating biomarkers could be the pleiotropic nature of *APOE* on lipid metabolism. Although *APOE* is expressed predominantly in the brain, it is also highly expressed in the adrenal gland, liver and kidney. *APOE* primarily plays a role in cholesterol metabolism as it targets lipoproteins and regulates cholesterol utilisation, potentially mediating the genetic associations in these analyses. However, if this is the case, it would be expected that other biomarkers involved in lipoprotein or cholesterol metabolism, such as LDL or ApoB, would also show a similar trend, yet the beta values were not as dramatically different in the third model for those biomarkers.

Although the present study controlled for the presence of *APOE* and CVD, this finding amongst fully adjusted models may still reflect the result of

comorbidities. Therefore, it is essential to consider whether the comorbidities differ across populations studied for exploratory analyses. For example, cardiovascular factors previously associated with blood biomarkers and dementia include hypercholesterolemia, hypertension, atherosclerosis, coronary heart disease, head injury, smoking, obesity, and diabetes. Populations with specific comorbidities or pathologies (e.g., cardiovascular conditions, *APOE* status or levels of AB positivity) may represent different risks. Therefore, it may be important to consider subtypes as this may allow refinement as biomarker profiles are likely to look different amongst different subtypes.

6.4.2 Biomarker and PGR by sex

When looking for biomarkers showing interactive effects, CRP, creatinine, LpA and oestradiol showed interactive effects for sex but were not associated with PGR of AD otherwise. When looking further by males and females separately, lower levels of creatinine, oestradiol and *CysC* were significantly associated with AD-PGR in females but not significantly higher for males, supporting the study of biomarker profiles differently among men and women. Levels of creatinine, *CysC* and oestradiol showed to be negatively associated for women only.

Although it is known that women are more at risk to develop AD, longevity does not explain this difference entirely. Stratification of individuals by sex in the present provides evidence that elevated risk of AD may present with sex specific patterns in the context of circulating blood biomarkers. However, the pathways behind these findings are not clear. One suggests the role of oestrogen related systems, including oestradiol, which has the highest affinity for intracellular receptors out of the three commonly considered estrogens and is used for hormone replacement therapies in postmenopausal women (Rabeya et al., 2021). Brain ageing occurs alongside endocrine ageing, where the role of such hormone changes should be considered. There are also reported changes in cellular and connectivity in the brain and cognition in tandem with decreasing oestradiol levels. However, evidence is conflicting regarding the association between oestrogen and dementia risk, where observational studies have shown evidence for and against protective effects (De Lange et al., 2020; Andrew et al., 2018). Some research suggests that oestradiol plays a protective role in regulating dendritic activity in hippocampal pyramidal cells in adulthood (Gould

et al., 1990). Numerous studies also have found that oestrogen may have protective effects against mitochondrial toxicity of Aß due to the production of antioxidants (Vina & Lloret., 2010; Andrew et al., 2018). These findings for oestrogen as a potential protective biomarker are consistent with this study which found that oestradiol was negatively associated with genetic risk of AD outside of *APOE* genotype.

In addition to the implication of oestrogen related systems, the relationship between Alzheimer's biomarkers and testosterone has been gaining attention, with recent evidence suggesting lower levels of testosterone are associated with high levels of p-tau (Alvarez de la Rosa et al., 2006). Although tau was not considered here, the present analysis found that testosterone showed a negative association, meaning lower levels of testosterone were associated with higher genetic risk. However, there were no sex interactions with PGR of AD and testosterone. That is, we found no evidence to suggest this sex difference is due to testosterone as a biomarker. This negative directionality, although not significant, may be due to a higher prevalence of AD amongst women compared to men (Andrew et al., 2018).

When looking at CRP, the directionality of association was in opposite directions for males and females. Other studies investigating sex differences in CRP concerning the development of AD have found that CRP tends to be higher in women (Duarte-Guterman et al., 2020; Laffoon et al., 2020). One study found that body fat was associated with CRP levels more in women than in men, suggesting adiposity may be a factor mediating this association (Khera et al., 2009). However, CRP levels can be highly variable due to their role in acute inflammation, and evidence is mixed for the directionality of association of CRP in AD. Some studies suggest that elevated levels of CRP in midlife can indicate a risk factor for cognitive decline and development of AD (Schmidt et al., 2002; Laurin et al., 2009), whereas studies looking at individuals in late life have found lower levels of CRP in individuals with mild to moderate AD compared to healthy controls, indicating a negative association (Yarchoan et al., 2020). Fernandes et al. (2020) suggested it may be helpful to use age when considering CRP to predict risk or disease progression of AD. Some findings suggest *APOE* e4 genotype modifies the association between CRP and onset of dementia, with

studies reporting lower levels of CRP in those carrying e4 genotypes (Ferguson et al., 2020; Fernandes et al., 2020).

Additionally, due to the immune system's complexity likely, there is not sufficient evidence in these analyses to comment on the associations between elevated genetic risk of AD and immune system biomarkers such as CRP. Clinical, observational, and animal studies suggest that the immune system is dysregulated in AD and is related to subsequent cognitive function and clinical status (Bettcher et al., 2021). However, this occurs in a non-linear manner which may vary amongst individuals according to immune system crosstalk and pathological stage of AD. That is innate and adaptive immune mechanisms and disease-specific pathogenesis such as the gut microbiome, amyloidosis, and peripheral cell infiltration (Bettcher et al., 2021).

Finally, although not significant, this study found that LpA was negatively associated with PGR for females and positively associated with PGR for males. In contrast, previous studies have suggested that sex hormones pose females at a higher risk of elevated LpA levels and related pathology than men. One study by Solfrizzi et al. (2002) found that: LpA serum concentrations were significantly associated with an increased risk for AD, independently of *APOE* genotypes and sex; however, this study found evidence to suggest there may be sex differences for LpA in the context of AD. It is also interesting to note that LpA was a biomarker of interest in chapter 4 where we found that higher blood LpA was associated with worse gMD and gFA. The present chapter (6) found LpA was negatively associated with PGR for females and positively associated with PGR for males. Together, these results suggest there may be sex specific effects of LpA that may play a role in WM structure that were not studied in detail as part of this thesis. However, although circulating blood biomarkers and cardiovascular risk factors may differ between men and women, extrapolating these findings to the cause of dementia may be more complex and a closer investigation into such differences may provide insights into precise associations.

6.4.3 Limitations and future research

The present study found several associations with blood biomarkers, genetic risk of AD, and sex interactions, which provide evidence for further biomarker

research. However, there are a number of limitations with these analyses that should be considered when considering the implications of the results. This study included only individuals of white European ancestry. There is evidence that genetic risk, circulating biomarkers, and relevant comorbidities vary among different populations, meaning this research cannot be generalised to a large proportion of individuals who will develop AD. This was a cross-sectional exploratory study in which further exploratory analyses, replication and longitudinal research is needed to understand how biomarkers can be utilised to predict future onset of disease.

A limitation posed by blood biomarker research is that some findings may reflect the difficulty of measuring brain-related biomarkers as it is not yet understood how accurately measures of circulating blood biomarkers can exhibit metabolic activity in the brain. This may be in part due to mediation of the BBB in which some measures may indirectly measure the integrity of the BBB (Elwood et al., 2017). This may be a confounding factor as permeability of the BBB is known to increase in AD, with some studies suggesting the direct role of *APOE* e4 (Halliday et al., 2016; Profaci et al., 2020). It is also challenging to determine if an altered concentration of a biomarker in the blood reflects activity in the brain or if it is secondary to comorbidity. Potential biomarkers may also present at low concentrations in blood after crossing the BBB, and it may be possible that standardisation or thresholds should be considered when investigating circulating biomarkers. It is also likely that genetic risk cannot explain all associations reported with blood biomarkers, and there are likely other conditions or environmental factors that need to be accounted for in relation to genetic risk.

6.4.4 Conclusion

This study found evidence for sex interactions with genetic risk of AD and circulating blood biomarkers in 502,536 individuals in the UK Biobank. We also found that *CysC* was associated with PGR-AD. However, there was little evidence for the association of other biomarkers in healthy adults.

Chapter 7: Thesis overview and discussion

7.1 Review of background and findings

Unhealthy brain ageing leading to cognitive decline in older age is a public health concern which is projected to become more prevalent, for which major risk factors include cardiometabolic conditions and genetics. The UK Biobank is a prospective cohort study that is well placed to examine these risk factors in the context of brain and address key gaps in our understanding of how they impact pathophysiology and subsequent brain health. This thesis aimed to address these gaps by examining two main research aims:

1. Contribute to the understanding of cardiovascular contributions to brain health, in particular multimorbidity.

2. To determine the role of genetic risk factors (for dementia) on the brain and physical health (indexed by biomarkers) in healthy adults.

7.1.1 Chapter-specific contributions

Cardiometabolic conditions are often comorbid and are established as contributors to worse brain health individually; however, multimorbidity outcomes are not well studied. Chapter 3 investigated associations between number and type of cardiometabolic comorbidity in relation to structural brain MRI. This chapter aimed to understand these associations on a larger scale than any previous study and examine differences in neuroanatomical substrates of cardiometabolic comorbidity, which are not currently established. Regression models examined eight mutually exclusive cardiometabolic groupings, and we found that individuals with no conditions presented with the healthiest brain metrics. On the other hand, individuals with two cardiometabolic conditions were associated with worse MRI measures for WMH, GM and whole-brain volume. Lack of trends with more specific subgroups suggests grouping by the number of conditions may be more beneficial to research than condition-specific subgroupings. This chapter established preliminary evidence to suggest that grouping individuals by the number of cardiometabolic comorbidities may be informative; however, an unexpected finding of this work was that individuals with most comorbidities (i.e., those reporting hypertension, type two diabetes, and coronary artery disease) did not present with the poorest brain MRI metrics. This may be due to a small sample size within the group or a consequence of

survival bias. The reason for this is not apparent, and it would be valuable to replicate or find conflicting evidence. Additionally, seven associations were no longer significantly associated with worse brain metrics when fully adjusted for covariates, implicating the contribution of external factors on brain health when considering cardiometabolic health.

Chapter 4 focused on a more specific and well-established biomarker for cardiovascular and cardiometabolic health; Lipoprotein A (LpA), by investigating differences between blood and genetically elevated LpA in brain MRI phenotypes. Lipoproteins are gaining interest in the role of brain ageing, as shown by an increase in lipidomic analyses and evidence of lipid pathways in cognitive ageing (Proitsi et al., 2017). Although cholesterol is the most studied lipid in the brain, lipids such as LpA may also be valuable to study due to their casual association with CVD. There is currently little literature examining the role of this risk factor in the context of brain ageing. Chapter 4 examined whether genetically elevated and circulating blood LpA is associated with brain MRI measures in healthy adults. We replicated previous associations when looking at blood lipid LpA and brain volumes; we found that blood LpA was associated with several measures of brain MRI, whereas genetically elevated LpA was associated with general factors of white matter microstructure: gMD and gFA. A possible explanation for the discrepancy between blood and genetically determined LpA associations with MRI may be due to genetically determined pathological processes not yet occurring, suggesting there may be a difference in blood and genetically elevated LpA that is not yet understood. Age of onset may be a crucial factor when considering the effects of genetic risk of LpA. It is also interesting to note that although both blood and genetic measures of LpA were associated with brain MRI, no associations were found for cognitive measures; however, it is important to remember for this cross-sectional analysis that the value of genetic risks scores for such outcomes can only be established over time. Further epidemiology studies of biomarkers may help us understand whether lipoproteins such as LpA ultimately affect cognitive ageing.

Chapter 5 used similar PGR scoring methods as chapter 4 but aimed to examine the association between polygenic risk for late-onset Alzheimer's Dementia (AD PGR) in relation to structural brain MRI and cognitive abilities. Chapter 5

questioned whether higher AD-PGR is associated with structural brain imaging and cognitive performance differences in a large sample of non-demented, generally healthy adults. We hypothesised that higher genetic risk would be associated with differences even in healthy adults. Summary statistics from an existing AD GWAS were used to create PGR scores for 32, 790 white European UK Biobank participants with LDpred. We tested for independent associations of polygenic risk (per SD), *APOE* e4 dose, and their interaction. *APOE* e4 dose was associated with several hippocampal volume measures and WMH volume as expected. We also found that AD-PGR was associated with worse hippocampal phenotypes. There were also associations of AD-PGR with worse cognitive performance in fluid intelligence and reaction time. *APOE* No evidence interaction between PGR and *APOE* dose was found when models were fully adjusted for social deprivation, *APOE* dose and smoking. These results suggested that PGR of AD is associated with smaller hippocampal volumes (both overall and subdivision specific) and higher WMH volumes in healthy adults without any evident cognitive impairments. These findings support a possibility of PGR scores supplementing *APOE* status in risk stratification of cognitive impairment/AD. However, we did not find any evidence to suggest that the genetic risk of AD may mediated by the most significant genetic risk factor, *APOE*.

Chapter 5 found evidence to suggest a PGR score of AD may indicate early signs of pathology in healthy individuals; as a result, chapter 6 aimed to use this genetic risk score and add further exploratory analyses looking at blood biomarkers and interactions by sex. Existing literature shows that differences in blood biomarker profile according to genetic risk of AD may help to elucidate the mechanisms at play when individuals are at elevated risk in midlife before any pathology is evident. Sex differences concerning genetic risk and blood biomarkers have been studied very little. Chapter 6 modelled associations between PGR of AD and 30 blood biomarkers and explored sex interactions with the genetic risk. We found that an increased polygenic risk of AD was associated with decreased levels of cystatin c (CysC) for both men and women, but particularly in women, which is consistent with existing literature. Sex differences were found for levels of creatinine, oestradiol and CysC. Lower levels of these were significantly associated with AD-PGR in females but not for males. The difference in biomarker profiles by sex provides further evidence to

an existing body that stratifying analyses by sex, particularly biomarker studies, will help bring to light underlying mechanisms.

7.1.2 Contributions to the literature

- Health outcomes of cardiometabolic comorbidity have not been well studied, despite studies indicating overlap between cardiometabolic conditions and decreased brain health. In chapter 3 of the present thesis, we compared eight mutually exclusive cardiometabolic subgroups, which is the first study to our knowledge to examine such conditions in the same cohort in relation to brain MRI. We found that grouping by the number of cardiometabolic conditions is of interest.

- There is evidence that the role of lipids is mechanistic to the ageing brain and cognitive decline in which investigation of candidate biomarkers may elucidate aetiology. Chapter 4 provided evidence to support the use of LpA as a risk factor for brain health, where there is currently little evidence for or against it. Chapter 4 showed that elevated LpA (both blood and genetically determined) is associated with poorer brain MRI measures in healthy adults, with blood LpA showing more associations than genetically elevated LpA. This is a novel contribution to the literature because there is little evidence to support or disprove the role LpA may play in brain and cognitive ageing.

- Chapter 5 found evidence that using AD-PGR demonstrated significant differences in non-demented brain structure in generally healthy individuals with a mean age of 64, suggesting PGR may be a helpful tool—in combination with other factors—for identifying individuals at risk of worse cognitive abilities and potentially accelerated decline. A major strength of this study is that previous studies testing associations between polygenic risk for late-onset Alzheimer's disease and brain magnetic resonance imaging measures have been limited by small samples and lack of controlling for confounders.

- Chapter 6 found evidence for sex differences in blood biomarkers according to genetic risk of AD PGR, which is a novel contribution to our knowledge of an elevated genetic risk of AD in a healthy population. We were able to provide evidence that sex stratification in biological research is valuable.

A primary role of using brain imaging within clinically relevant research has been to aid in diagnosis and to monitor disease progression. It is less common for brain MRI metrics to provide predictive indicators or help to stratify individuals. This thesis, however, has been able to make use of large sample sizes of healthy adults and examine imaging markers of early disease to offer preliminary evidence for predictive markers in relation to other risk factors, i.e., genetic or cardiometabolic. This allows for insights into cognitive ageing at its earliest stages, which could potentially provide initial indicators of disease mechanisms.

Although we did not draw conflicting conclusions to the consensus within any chapters, there were some discrepancies between variables when looking to the wider literature in chapter 3. It was difficult to compare findings with similar papers due to the nature of the independent variables used, with other studies using more specific variables. For example, Beck et al. (2021), Kolbeinsson (2020) and Vergoossen et al. (2020) were comparable studies that examined cardiometabolic related risk factors and conditions in relation to brain MRI. Beck found the most significant contributors to faster brain ageing were high systolic blood pressure, waist-hip ratio and smoking. Vergoossen found the most significant risks were hyperglycaemia, physical activity, obesity, and hypertension. Kolbeinsson found evidence to implicate hypertension in larger brain age gaps. Although lack of comparable studies and measures do not explain the inconsistency of results in chapter 3, it calls attention to the gap in research for multimorbidity.

7.1.3 Implications

This thesis emphasises the importance of both genetic and cardiometabolic contributions to brain and cognitive outcomes. There is evidence within all chapters that contribute to understanding associations between several complex

phenotypes better. Chapter 6 emphasised the importance of stratification of genetic research in this area by sex, where we suggested such stratification by should currently be routine. Chapter 3 also provided evidence that stratification by number of cardiometabolic conditions aided in elucidating significant differences in brain MRI. However, genetic and cardiometabolic factors are highly dynamic and unravelling mechanisms by more than one biological level may be a better aim than taking a reductionist approach to stratify individuals by two categories. This is easier to research within cohort and population studies and implementing multiple levels of stratification by individual risk factors may help identify the largest risks in relation to the most significant contributors. This thesis studied this at a population level instead of an individual level, meaning findings are better suited to apply within clinical and public health settings; for example, there is scope to educate those at risk of the importance of cardiometabolic health with consideration of genetics.

As reported in chapters 5 and 6, PGR of AD was associated with features of abnormal brain ageing and biomarker differences. PGR scores alongside clinical data may ultimately help to better understand specific aspects of an individual's brain health and pathology and potentially predict clinical presentations. However, PGR scores for AD and LpA in this thesis were used as research tools to understand the degrees of correlation. It is important to remember that different statistical approaches can be used to develop PGR scores, in which all models are then rigorously validated for their value in predicting a disease state or trait. PGR scoring is not standardised, and many factors can affect the models, including data used, different statistical approaches to selection and weighting of relevant variants, and the process of interpreting the output can also influence findings. Early studies of genetic variants were shown not to replicate consistently (Mufano et al., 2006), and variations in conducting GWAS were thought to contribute to this. However, reporting standards are becoming more common as the field has grown (Wand et al., 2021). Moreover, clinical practice or applications should not change based on exploratory research as standardisation criteria are required.

7.2 Thesis limitations

7.2.1 Selection bias

The UK Biobank assessment centre recruitment took place over 22 centres in which the breakdown of each centre can be found here: https://biobank.ctsu.ox.ac.uk/showcase/exinfo.cgi?src=UKB_centres_map. There have been several papers assessing the validity and reliability of the population used within UK Biobank that has concluded there is a healthy volunteer sampling bias in which participants have less cancer incidence, drink less alcohol, have fewer self-reported diseases, and live in less socioeconomically deprived areas (Fry et al., 2017; Tyrell et al., 2021; Lyall et al., 2021). The lack of generalisability and unequal collection of data may be limiting to the research in this thesis. One such way is due to restricted range of measures for cognition due to a healthy cohort, leading to potential underestimates of association. It is possible that valuable indicators of brain health or biomarkers could have been overlooked due to unmeasurable or unmeasured biases due to a healthy sample (e.g., deprivation, gender, ethnicity etc.) that the findings of this thesis will perpetuate. Potential consequences include misdiagnosis, underdiagnosis, and increasing healthcare disparities.

7.2.2 Lack of diversity

All analyses using genetic data in this thesis were carried out in healthy middle-aged White European participants. This sample may be unrepresentative of those with the highest genetic risk, as it is a healthy cohort. It is possible that thesis may not have studied those with the most increased genetic risks and, therefore, those individuals who will go on to present with phenotypes of interest. However, some research on the UK Biobank cohort shows that although risk factors levels were favourable towards UK Biobank compared to nationwide registries, levels of associations would suggest findings are still generalisable (Batty et al., 2019).

7.2.3 Collider bias

The problem of potential confounders in the UK Biobank was challenging to navigate in this thesis as it was difficult to account for all confounding variables

which may have led to residual or unmeasured confounding. Additionally, when the underlying sampling model is biased, for example the UK Biobank is likely a healthy cohort in comparison to the general population, large sample size may magnify the bias the population presents. In this thesis this could mean that results where confounders such as smoking were controlled for may not accurately represent the association for other individuals in terms of the independent variable, the dependent variable and the role of smoking. We chose to regress out confounders from analyses within our models; we did this with age, sex, BMI, CVD medication, Townsend deprivation score, education, smoking, and included principal components where genetic data was used. There are strengths and weaknesses to this approach we used as it could be argued that age or sex, for example, should not be considered a confounder but a variable of interest. For example, some studies using these variables as predictors by carrying out sex-stratified analyses and age-specific effects have found this method valuable in understanding individual differences (Guerreiro & Van Gerven, 2011; Salami et al., 2012; Lamar et al., 2020). A more critical approach may have been to carry out separate association analyses, e.g., for age groups to examine whether age should be treated as a confounder or not. Lack of further examination of these factors is a limitation of this work as many associations in this thesis were no longer significant when models were fully adjusted using demographic data such as education, Townsend or smoking. Although beyond the scope of this thesis, accounting for other environmental influences may have also helped to explain associations we found. Due to the way we chose to account for potential confounders in this thesis, it was not possible to know the extent to which variance in outcome variables such as brain imaging was due to confounders or genuine variance. This has been overlooked but highlights the importance of considering these factors closely.

Some variables we included as confounders are nuanced and related to pathology in complex ways such as age, sex, sociodemographic status. One example of relevance to this work is age as a phenotype. Some research has shown that biological age measures (such as brain age gap) are better at predicting ageing outcomes than chronological age. Hou et al. (2019) described nine biological hallmarks of ageing: genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, mitochondrial dysfunction, cellular

senescence, deregulated nutrient sensing, stem cell exhaustion, and altered intercellular communication. The paper suggested that these central mechanisms of biological ageing should be considered more closely in neurodegenerative diseases and suggest that age is not necessarily linear in the context of the brain. One study using UK Biobank calculated biological (age-adjusted) phenotypes instead of chronological age. They were calculated as functions of chronological age, albumin, creatinine, C- reactive protein (CRP), alkaline phosphatase, glucose, lymphocyte percentage, mean corpuscular volume, red blood cell distribution width, white blood cell count, glycated haemoglobin, systolic blood pressure, and total cholesterol. These biological age phenotypes have been shown to be robust predictors of ageing outcomes (Levine, 2013; Levine et al., 2018). This suggests ageing may not be most accurately captured in chronological age alone and raises the question of whether we can better define age as a risk factor for cognitive and brain health.

Alongside age, there are several additional confounders that these analyses could not consider that have consistently been implicated in various aspects of health, including cognitive health. Objective physical activity and objective sleep measures could have been included in this thesis; however, this data was only collected in a subset of 100k UK Biobank participants (Doherty et al., 2017; Gill, 2020). Sociobehavioural proxies are known to play a role in explaining cognitive reserve and cognitive resilience, which are central to avoiding cognitive impairment and can help foster cognitive health (Radanovic, 2020). Such variables available in the UK Biobank that could have been considered in this thesis include sleep, nutrition, depression, parental socioeconomic position, air pollution exposure, and social activity (Cullen et al., 2018; 2019; Lyall et al., 2019). Psychosocial health is also a factor known to influence the brain and cognitive health and was not considered here; one study by Cullen et al. (2019) found in the UK Biobank that cognitive impairment was associated with mood disorders and psychotropic medication use. It is also known that aspects of the blood-brain barrier are controlled by the circadian clock, which is strongly associated with mood dysregulation, further implicating the role of psychosocial factors. Research has also shown that allostatic load, wearing due to exposure to chronic stress, may contribute to the disproportionate ageing and health outcomes (Higgins et al., 2019).

7.2.4 Genetic risk scoring

The PGR scores in this thesis may be useful for association analyses, more so than for clinical practice or predicting disease, as it is not likely that PGR scores will be used as standalone tools but combined with other risk estimates. It may be in some cases that using fewer variants may improve accuracy as with large numbers of variants; for example, the ~6 million in chapter 5, it is not possible to understand how each SNP relates to the outcome phenotype. This can be problematic as there are likely to be some variant associations included within this thesis resulting from spurious associations. If the aim of using such tools are more refined diagnoses and improved precision, identifying polymorphic variants with clear functional effects on complex traits may be more practical than considering many variants which may not play a role in the molecular basis of the complex trait of interest. This approach seems closer to implementation and is known as 'partitioned' or 'partial' polygenic scores (Barroso & McCarthy, 2019). Using such a polygenic score model alongside non-genetic risk factors will likely provide a combined risk estimate in the context of clinical applications. However, it is important to remember that the number of genetic variants that influence a phenotype and the size of their effects on the phenotype vary depending on the trait. For this reason, the optimal number of variants to consider for a PGR score depends on the underlying genetic architecture and will also vary.

APOE genotype was considered as either 0/1/2 depending on number of e4 alleles. This was because we were interested in whether an individual was likely to have a risk posed by the *APOE* gene or not. There are limitations to this as approach as opposed to using all *APOE* haplotypes e.g. e2/e2, e2/e3 etc. Limitations to not considering refined haplotypes of *APOE* include not fully capturing differences in genetic risk between variations of *APOE* genotype. The e4 allele count is commonly included due to the established genetic risk it poses for AD, MCI and cognitive decline (Stocker et al., 2018; Abondio et al., 2019; Shi et al., 2019; Zhao et al., 2020). Although there is evidence that consideration of further haplotypes can be informative, chapters 4 and 5 of this thesis only consider the number of e4 alleles for *APOE* risk. This was because *APOE* ε4 allele remains the largest genetic risk factor for late-onset AD and the analyses within

these chapters aimed to study whether genetics risks may differ according to number of e4 alleles.

7.2.5 Cross-sectional data

It is also not possible to establish causality from differences within cross-sectional structural MRI measures; one study found that deep WMH and periventricular WMH, which are strongly associated with abnormal ageing and dementia, have different genetic underpinnings, suggesting that even well-established MRI phenotypes still require closer examination when considering causality in this area (Armstrong et al., 2020). It has been challenging within the work of this thesis to draw justified conclusions about brain imaging differences between individuals, as lower brain volume associations cannot be attributed to pathology; for example, a paper by Wheater et al. (2021) found that birth weight was more strongly associated with brain tissue reserve in later life than age-related structural features such as tissue atrophy or WMH was. However, more deeply phenotyped measures of brain structure and features may help with such challenges to differentiate normal and abnormal changes and attribute pathological causes. Analyses in this thesis involving MRI imaging variables were carried out in a healthy cohort for which there may not have been sufficient age-related for phenotypes to present. However, it is notable that we found associations with *APOE* genotype that would be expected and that are consistent with the literature, indicating the findings of the imaging analyses may still be informative and can contribute to the literature on cardiometabolic and genetic associations.

Chapters 4 and 6 studied circulating blood biomarkers that were taken at baseline. However, covariates included within analyses and imaging data were taken at different instances and there may have been an 8 year difference between collection of different variables which were analysed within the same model. This could lead to inaccuracies of the data as biological biochemistry is likely to change over time. The lack of cross-sectional data for these chapters raises important considerations around validity of the cross sectional associations and inferences made. However, A study by Trinder et al., (2020) found that the *LpA* biomarker within the UK Biobank remained stable which they

defined as having minimal influence from age, sex, genetic factors outside the *LPA* gene, environmental factors, or currently available medicines.

7.2.6. Multiple testing

This thesis did not correct for multiple testing which has a number of limitations. The most significant limitation to the lack of multiple testing correction is the possibility of making false inferences due to false positive associations reported with simultaneous investigation of more than one research question. Adjusting statistical inference usually is done to help avoid declaring associations when there are none, and in turn help in making valid scientific conclusions. A number of procedures have been developed to deal with controlling for appropriate error rates, such as the Bonferroni correction, but there is continuing controversy around how to best balance false positive and false negative results. Williams & Haines (2011) argue that correcting for multiple testing may be overly conservative and result in studies with inflated false negative results. The paper argues that a statistical significance threshold may not be the most accurate approach to determining which findings are by true effects and which are not. A main research objective of this thesis focused on carrying out association analyses of phenotypes with consideration of confounding variables to formulate more specific research hypotheses for factors influencing phenotypes of interest. Validation, replication and confirmatory analyses of these findings should be carried out where type 1 errors are more easily identified in subsequent studies. However, it is important to note that if analyses had been adjusted for multiple testing, there are multiple findings in each chapter that would not have survived correction for significance. This would consequently change the conclusions and implications of the findings within this thesis.

7.3 Future research

7.3.1 Image-Derived Phenotypes

This thesis used the UK Biobank image-derived phenotypes (IDPs) for which processing and quality control procedures are described in: Chapter 2: Methods.

These IDPs have been examined within the literature in which there has been discussion around the validity and reliability of IDPs as phenotypes. The two main limitations of the IDPs used for this thesis are that a broader range of neuroimaging parameters could have been used and would be more informative. Additionally, the IDPs could not capture relevant brain features to cardiometabolic and cognitive health. Investigations of CVD and neurovascular dysfunction in dementias have primarily focused on structural brain imaging similarly to this thesis; however, multi-modal imaging can reveal more information about mechanisms of interest such as rates of blood flow and functional connectivity. Specific measures to cardiometabolic health include periventricular spaces (Wardlaw et al., 2020) and gyrification, which has been shown to decrease with age, including precentral, temporal and frontal areas (Gennatas et al., 2017).

7.3.2 Multi-omics

Omics technologies - such as genomics, transcriptomics, proteomics, and metabolomics - in addition to evolving imaging techniques - may allow us to understand a larger number of molecular features in relation to phenotypes of interest. This may provide the potential to consider more specific and relevant environmental factors. Capturing environmental and biological factors of a phenotype comprehensively with “multi-omics” may give researchers the ability to map complex interactions and carry out assessments in a precise way, leading to understanding of causal pathways in cognitive ageing. There is potential for omics research within cohorts such as UK Biobank to allow for longitudinal analysis, providing a more detailed picture of the biology underlying subtypes and contribute to precision medicine approaches. It is understood that hallmark biological characteristics of ageing include genomic instability, telomere attrition, epigenetic alterations, protein loss and mitochondrial dysfunction (Melzer et al., 2020) in which all other age-related phenotypes are proxies of this. With better omics technologies, these hallmarks can be measured directly in human genetic association studies and potentially reveal more about missing heritability. Of specific interest to this area is better understanding of the ways by which *APOE* e4 modulates biological pathways. For example, it is not yet understood whether *APOE* e4 influences the rate of ageing mechanisms or

neurodegeneration (Henson et al., 2020). Multi-omics could also be extended to address gene and environment interactions of complex gaps in our knowledge. For example, Andrew & Tierney (2018) outline the drivers of differences between men and women in dementia and possibly other age-related cognitive decline, including longevity, biological differences, differences in cognitive performance, and gendered social roles and opportunities. Another method to potentially investigate gene and environment interactions by omics is to examine methylation patterns associated with lifestyle traits (Gadd et al., 2021).

Another potential area for investigation is a better understanding of imaging through the application of omics technologies; one example where this has been carried out was in a paper by Elliot et al. (2019), who examined the heritability of UK Biobank's IDPs to understand the genetic architecture of the brain. With the use of multi-omics, further similar research could be carried out to uncover causal pathways that link genetic variants to IDPs and other brain imaging metrics to a range of psychiatric and ageing disorders. One current method used to approach such investigations is Mendelian Randomization (MR). MR is a statistical method that utilises genetic variants that have a well-understood function to act as an instrumental variable to establish casual relationships. Genetic variants are used as proxies for an exposure where applications and developments within MR methodology show the use of MR to inform drug development and for phenome-wide studies (Schmidt et al., 2017; Li et al., 2019; Larsson et al., 2020; Zheng et al., 2020). Mufano & Davey-Smith (2018) argue that both replication and triangulation of findings is critical, which they describe as the strategic use of more than one methodology or approach to address a research question or provide evidence for a relationship. Each methodology has different and unrelated sources of bias; for example, within MR, many assumptions must be met. Notably, all confounders must be measured and fully controlled for, and primary sources of bias for this method are typically due to linkage disequilibrium, pleiotropy, population stratification and canalisation.

7.3.3 Sex stratification

Most complex traits, including cognitive ageing, are influenced by sex differences to a greater or lesser extent, which ultimately affects outcomes such as disease progression, presentation of symptoms, or the age of onset – phenotypes that are not typically studied in the context of sex. It is understood that brain ageing, cardiometabolic health and relevant mechanisms such as lipid metabolism are directly modulated by estrogen and testosterone. One study examined 530 traits within the UK Biobank and found that 71 traits presented significantly different heritability estimates between the sexes (Bernabeu et al., 2021). Studies elucidating sex differences in cardiometabolic disorders have found that women have a higher cardiometabolic risk factor burden than men, where female sociocultural factors such as caregiving responsibilities were more associated with cardiometabolic risk factors than sex (Gerdts & Regitz-Zagrosek, 2019). Additionally, when compared to men of the same age, the prevalence of metabolic syndrome is lower in premenopausal women but higher in post-menopausal women. The literature strongly supports that sex and gender play central roles in disease pathologies and mechanisms, where the current approach to studying sex differences is an area for future research to improve upon in this field. Dichotomising sex as a variable may not help to understand disease aetiologies in their full complexity; Khramtsova et al. (2019) suggest that the term sexual dimorphism has been misused within research to describe two distinct forms of a phenotype where many sex differences exist on a spectrum. Instead, it may be beneficial to phenotype sex and gender more specifically to understand these differences.

7.3.4 Increased diversity

Another current barrier to progress in this area is that research cohorts such as the UK Biobank often lack representativeness, preventing inequalities from adequately being examined. The findings of this thesis and a large proportion of genetic research in brain and cardiometabolic health as it stands are not representative enough to be applied to individuals who do not have white European ancestry. Genetics has contributed significantly to our understanding of disease processes; however, complex diseases often result from many

interactions between biology and environment (including societal factors), both of which are understudied in many populations. Studies on cardiovascular disease epidemiology suggest that socioeconomic position and education play a role in developing modifiable risk factors for cardiovascular disease such as BMI, smoking, and blood pressure (Smith et al., 1997). In the literature, the socioeconomic position is typically measured at an individual level, e.g., educational attainment, and at the population level, e.g. index of deprivation, where socioeconomic status implies status determined by societal norms. However, it is more likely to result from material resources such as income (Krieger, 2002). Krieger states that "health disparities, within and between countries, that are judged to be unfair, unjust, avoidable, and unnecessary…systematically burden populations rendered vulnerable by underlying social structures and political, economic, and legal institutions" (Krieger, 2001 pp 72). The first Marmot review examined health inequalities in England and found that if mortality rates were the same between the least and most deprived individuals, approximately 2 million extra years of life could have been lived (Marmot, 2010). It also provided evidence to suggest that standardising healthcare access would help address many health inequalities. Such modifiable risk factors that may cause health inequalities are challenging to consider in models as covariates, and better consideration needs to be given as to how genetics and the environment work together to contribute to disease risk. This is particularly relevant as one of the primary sources of bias is confounding, where both an exposure and an outcome share a common cause. For example, adjusting for years of education attained cannot account for differences in experience where there are historic and systemic inequalities (e.g. along ethnic, gender, class lines). Genetic research may not be an objective way to address this, as an overemphasis on genetics as an explanatory factor may contribute to disparities or reinforce stereotyping, perpetuating disparities and overlooking environmental contributions.

7.3.5 More representative data

It is important to note that these factors cannot be considered in isolation and should be accompanied by reporting other health determinants and sources of inequality. For example, the intersectionality of race and ethnicity for which there is comprehensive literature showing the effects of systemic racism within

health care and research. Focusing solely on differences in racial categories may be an area of improvement within the current field, as these are socially constructed classifications that are not treated as such within the research (Flanagin et al., 2021). It is crucial to consider the downstream effects of exclusion or oversimplification as accompanying assumptions may be built into the concepts and methods used, creating the potential for further discrimination. One example of this is within unrepresentative research cohorts, e.g., Framingham Heart Study or UK Biobank, where research findings are overgeneralised to those outside the study sample, creating the potential to perpetuate a disproportionate burden of disease. This is an area for research to improve upon by carrying out inclusive research. It can also be addressed within studies that consider study design by matching or restricting or causal models where temporality between exposures and outcomes can be better accounted for. There are advantages of including underrepresented individuals to investigate genetic variation. For example, population stratification is typically adjusted for using principal genetic components as this thesis did; however, software such as Tractor allows for the inclusion of admixed individuals in GWAS. The inclusion of admix populations is good for the further discovery of contributing or causal variants, in which the use of ancestry-specific SNP effect sizes has been suggested for the development of PGR scores (Atkinson et al., 2021). This approach also has the potential to shed light on missing heritability.

7.4 Summary

This PhD thesis investigated the effects of two major risk factors for brain health: cardiometabolic disease and genetics within a large cohort. As part of this work, chapters examined the genetics of Alzheimer's disease dementia in the context of overall health (i.e., biomarkers) and brain health, brain correlates of a well-established CVD risk factor (*LpA*), and cardiometabolic comorbidity. This thesis also considered demographic variables when looking at conditions, specific biomarkers and genetic risks for both cardiovascular disease and dementia. The overall results of this thesis found consistent evidence that measures of cardiometabolic health were associated with worse brain health measures compared to healthy individuals, including microstructural integrity and hippocampal volumes. The findings of this thesis have contributed to understanding the role of cardiometabolic health in brain health. Specifically,

chapter 4 found that blood *LpA* and PGR of elevated *LpA* was associated with worse brain MRI phenotypes. Chapter 5 found that genetic risk of AD, which implicated cardiometabolic variants, was associated with worse brain MRI in healthy adults before any cognitive symptoms were evident. Chapter 6 found cystatin C was associated with elevated genetic risk for AD and found evidence for the mediating role of sex in genetic risk of AD and circulating blood biomarkers. Chapter 7 summarised the key findings of the research from this thesis and outlined the key strengths and limitations of analyses. The limitations section considered potential analyses and variables that were not considered in this thesis, and areas for future work have been discussed.

References

Abondio, P., Sazzini, M., Garagnani, P., Boattini, A., Monti, D., Franceschi, C., … Giuliani, C. (2019). The genetic variability of APOE in different human populations and its implications for longevity. *Genes*, *10*(3). https://doi.org/10.3390/genes10030222

Abraham, G., Malik, R., Yonova-Doing, E., Salim, A., Wang, T., Danesh, J., … Dichgans, M. (2019). Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nature Communications*, *10*(1), 1–10. https://doi.org/10.1038/s41467-019-13848-1

Alvarez-De-La-Rosa, M., Silva, I., Nilsen, J., Pérez, M. M., García-Segura, L. M., Ávila, J., & Naftolin, F. (2005). Estradiol prevents neural tau hyperphosphorylation characteristic of Alzheimer's disease. *Annals of the New York Academy of Sciences*, *1052*, 210–224. https://doi.org/10.1196/annals.1347.016

Andrew, M. K., & Tierney, M. C. (2018). The puzzle of sex, gender and Alzheimer's disease: Why are women more often affected than men? *Women's Health*, *14*, 174550651881799. https://doi.org/10.1177/1745506518817995

Armstrong, N. J., Mather, K. A., Mather, K. A., Sargurupremraj, M., Knol, M. J., Malik, R., … Nyquist, P. A. (2020). Common Genetic Variation Indicates Separate Causes for Periventricular and Deep White Matter Hyperintensities. *Stroke*, (July), 2111–2121. https://doi.org/10.1161/STROKEAHA.119.027544

Arnoldussen, I. A. C., Gustafson, D. R., Leijsen, E. M. C., de Leeuw, F. E., & Kiliaan, A. J. (2019). Adiposity is related to cerebrovascular and brain volumetry outcomes in the RUN DMC study. *Neurology*, *93*(9), e864–e878. https://doi.org/10.1212/WNL.0000000000008002

Atkinson, E. G., Maihofer, A. X., Kanai, M., & Martin, A. R. (2021). Tractor uses local ancestry to enable inclusion of admixed individuals into GWAS and boost power Elizabeth. *Nat Genet.*, *176*(1), 139–148. https://doi.org/10.1038/s41588-020-00766-y.Tractor

Attems, J., & Jellinger, K. A. (2014). The overlap between vascular disease and Alzheimer's disease - lessons from pathology. *BMC Medicine*, *12*(1), 1–12. https://doi.org/10.1186/s12916-014-0206-2

Backholer, K., Peters, S. A. E., Bots, S. H., Peeters, A., Huxley, R. R., & Woodward, M. (2017). Sex differences in the relationship between socioeconomic status and cardiovascular disease: A systematic review and meta-analysis. *Journal of Epidemiology and Community Health*, *71*(6), 550–557. https://doi.org/10.1136/jech-2016-207890

Bangen, K. J., Preis, S. R., Delano-Wood, L., Wolf, P. A., Libon, D. J., Bondi, M. W., … Brickman, A. M. (2018). Baseline white matter hyperintensities and hippocampal volume are associated with conversion from normal cognition to mild cognitive impairment in the Framingham Offspring Study. *Alzheimer Dis Assoc Disord*, 50–56. https://doi.org/10.1097/WAD.0000000000000215

Barroso, I., & McCarthy, M. I. (2019). The genetic basis of metabolic disease. *Cell*, *177*(1), 146–161. https://doi.org/10.1016/j.cell.2019.02.024.The

Batty, G. D., Gale, C. R., Kivimäki, M., Deary, I. J., & Bell, S. (2019). Comparison of risk factor associations in UK Biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *The BMJ*, *368*, 1–8. https://doi.org/10.1136/bmj.m131

Beck, D., de Lange, A. M., Maximov, I. I., Richard, G., Andreassen, O. A., Nordvik, J. E., & Westlye, L. T. (2021). White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction. *NeuroImage*, *224*(July 2020), 117441. https://doi.org/10.1016/j.neuroimage.2020.117441

Bernabeu, E., Canela-Xandri, O., Rawlik, K., Talenti, A., Prendergast, J., & Tenesa, A. (2021). Sex differences in genetic architecture in the UK Biobank. *Nature Genetics*, *53*(9), 1283–1289. https://doi.org/10.1038/s41588-021-00912-0

Bettcher, B. M., Tansey, M. G., Dorothée, G., & Heneka, M. T. (2021). Peripheral and central immune system crosstalk in Alzheimer disease — a research prospectus. *Nature Reviews Neurology*, *17*(11), 689–701. https://doi.org/10.1038/s41582-021-00549-x

Biesbroek, J. M., Weaver, N. A., & Biessels, G. J. (2017). Lesion location and cognitive impact of cerebral small vessel disease. *Clinical Science*, *131*(8), 715–728. https://doi.org/10.1042/CS20160452

Birgittawerner, K., Elmståhl, S., Christensson, A., & Pihlsgård, M. (2014). Male sex and vascular risk factors affect cystatin C-derived renal function in older people without diabetes or overt vascular disease. *Age and Ageing*, *43*(3), 411–417. https://doi.org/10.1093/ageing/aft191

Blennow, H. K., Zetterberg, Blennow, K., Blennow, K., Hampel, H., Weiner, M., & Zetterberg, H. (2010). Cerebrospinal fluid and plasma biomarkers in Alzheimer disease. *Nature Publishing Group*, *6*(16104), 131–144. https://doi.org/10.1038/nrneurol.2010.4

Blennow, K. (2017). A Review of Fluid Biomarkers for Alzheimer's Disease: Moving from CSF to Blood. *Neurology Therapy*, 15–24.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Brinton, R. D., Yao, J., Yin, F., Mack, W. J., & Cadenas, E. (2015). Perimenopause as a neurological transition state. *Nature Reviews Endocrinology*, *11*(7), 393–405. https://doi.org/10.1038/nrendo.2015.82

Casanova, R., Varma, S., Simpson, B., Kim, M., An, Y., Saldana, S., … Thambisetty, M. (2016). Blood metabolite markers of preclinical Alzheimer's disease in two longitudinally followed cohorts of older individuals. *Alzheimer's and Dementia*, *12*(7), 815–822. https://doi.org/10.1016/j.jalz.2015.12.008

Cegla, J., Neely, R. D. G., France, M., Ferns, G., Byrne, C. D., Halcox, J., … Soran, H. (2019). HEART UK consensus statement on Lipoprotein(a): A call to action. *Atherosclerosis*, *291*(February), 62–70. https://doi.org/10.1016/j.atherosclerosis.2019.10.011

Chew, H., Solomon, V. A., & Fonteh, A. N. (2020). Involvement of Lipids in Alzheimer's Disease Pathology and Potential Therapies. *Frontiers in Physiology*, *11*(June), 1–28. https://doi.org/10.3389/fphys.2020.00598

Cohen, R. A., Poppas, A., Forman, D. E., Hoth, K. F., Haley, A. P., Gunstad, J., … Gerhard-Herman, M. (2009). Vascular and cognitive functions associated with cardiovascular disease in the elderly. *Journal of Clinical and Experimental Neuropsychology*, *31*(1), 96–110. https://doi.org/10.1080/13803390802014594

Cole, J. H. (2020). Multimodality neuroimaging brain-age in UK biobank: relationship to biomedical, lifestyle, and cognitive factors. *Neurobiology of Aging*, *92*, 34–42. https://doi.org/10.1016/j.neurobiolaging.2020.03.014

Congdon., E., & Sigurdsson., E. (2018). Tau-targeting therapies for Alzheimer disease. *Nature Reviews Neurology*, *176*(10), 139–148. https://doi.org/10.1016/j.physbeh.2017.03.040

Counts, S. E., Ikonomovic, M. D., Mercado, N., Vega, I. E., & Mufson, E. J. (2016). Biomarkers for the Early Detection and Progression of Alzheimer's Disease. *Neurotherapeutics*, 1–19. https://doi.org/10.1007/s13311-016-0481-z

Cox, S. R., Lyall, D. M., Ritchie, S. J., Bastin, M. E., Harris, M. A., Buchanan, C. R., … Deary, I. J. (2019). Associations between vascular risk factors and brain MRI indices in UK Biobank. *European Heart Journal*, *40*(28), 2290–2299. https://doi.org/10.1093/eurheartj/ehz100

Crean, S., Ward, A., Mercaldi, C. J., Collins, J. M., Cook, M. N., Baker, N. L., & Arrighi, H. M. (2011). Apolipoprotein E ε4 prevalence in Alzheimer's disease patients varies across global populations: A systematic literature review and meta-analysis. *Dementia and Geriatric Cognitive Disorders*, *31*(1), 20–30. https://doi.org/10.1159/000321984

Cullen, B., Newby, D., Lee, D., Lyall, D. M., Nevado-Holgado, A. J., Evans, J. J., … Cavanagh, J. (2018). Cross-sectional and longitudinal analyses of outdoor air pollution exposure and cognitive function in UK Biobank. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-30568-6

Cullen, B., Smith, D. J., Deary, I. J., Pell, J. P., Keyes, K. M., & Evans, J. J. (2019). Understanding cognitive impairment in mood disorders: Mediation analyses in the UK Biobank cohort. *British Journal of Psychiatry*, *215*(5), 683–690. https://doi.org/10.1192/bjp.2019.188

Dani, M., Brooks, D. J., & Edison, P. (2017). Suspected non Alzheimer's pathology – Is it non-Alzheimer's or non-amyloid? *Ageing Research Reviews*, *36*, 20–31. https://doi.org/10.1016/j.arr.2017.02.003

Dara L. Dickstein, Walsh, J., Brautigam, H., Jr, S. D. S., Gandy, S., & Hof, P. R. (2010). Role of Vascular Risk Factors and Vascular Dysfunction in Alzheimer's Disease. *Mt Sinai J Med*, 15–21. https://doi.org/10.1002/MSJ

Davies, G., Lam, M., Harris, S. E., Trampush, J. W., Luciano, M., Hill, W. D., … Deary, I. J. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, *9*(1), 1–16. https://doi.org/10.1038/s41467-018-04362-x

De Lange, A. M. G., Barth, C., Kaufmann, T., Maximov, I. I., van der Meer, D., Agartz, I., & Westlye, L. T. (2020). Women's brain aging: Effects of sex-hormone exposure, pregnancies, and genetic risk for Alzheimer's disease. *Human Brain Mapping*, *41*(18), 5141–5150. https://doi.org/10.1002/hbm.25180

De Leeuw, F. A., Peeters, C. F. W., Kester, M. I., Harms, A. C., Struys, E. A., Hankemeier, T., … Teunissen, C. E. (2017). Blood-based metabolic signatures in Alzheimer's disease. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring*, *8*, 196–207. https://doi.org/10.1016/j.dadm.2017.07.006

Deary, I. J. (2013). Intelligence. *Current Biology*, *23*(16), 673–676. https://doi.org/10.1016/j.cub.2013.07.021

Deary, I. J., Harris, S. E., & Hill, W. D. (2019). What genome-wide association studies reveal about the association between intelligence and physical health, illness, and mortality. *Current Opinion in Psychology*, *27*, 6–12. https://doi.org/10.1016/j.copsyc.2018.07.005

Debette, S., & Markus, H. S. (2010). The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: systematic review and meta-analysis. *Bmj*, *341*(jul26 1), c3666–c3666. https://doi.org/10.1136/bmj.c3666

Debette, S., Seshadri, S., Beiser, A., Au, R., Himali, J. J., Palumbo, C., … DeCarli, C. (2011). Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology*, *77*(5), 461–468. https://doi.org/10.1212/WNL.0b013e318227b227

Deng, A., Irizarry, M. C., Nitsch, R. M., Growdon, J. H., & Rebeck, G. W. (2001). Elevation of cystatin C in susceptible neurons in Alzheimer's disease. *American Journal of Pathology*, *159*(3), 1061–1068. https://doi.org/10.1016/S0002-9440(10)61781-6

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M. H., … Wareham, N. J. (2017). Large scale population assessment of physical activity using wrist worn accelerometers: The UK biobank study. *PLoS ONE*, *12*(2), 1–14. https://doi.org/10.1371/journal.pone.0169649

Duarte-Guterman. (2020). Alzheimer's disease neuropathology, inflammation and epigenetic age: do sex and APOE matter? *Preprint*.

Dudbridge, F. (2013). Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genetics*, *9*(3). https://doi.org/10.1371/journal.pgen.1003348

Dufouil, C., De Kersaint-Gilly, A., Besançon, V., Levy, C., Auffray, E., Brunnereau, L., … Tzourio, C. (2001). Longitudinal study of blood pressure and white matter hyperintensities: The EVA MRI cohort. *Neurology*, *56*(7), 921–926. https://doi.org/10.1212/WNL.56.7.921

E. Gould, C.S. Woolley, M. Frankfurt, B. S. M. (1990). Gonadal Steroids Regulate Dendritic Spine Density in Hippocampal Pyramidal Cells in Adulthood. *The Journal of Neuroscience*, (April).

Elliott, L., Sharp, K., Alfaro-Almagro, F., Douaud, G., Miller, K., Marchini, J., & Smith, S. (2018). Genome-wide association studies of brain structure and function in the UK Biobank. *Nature*. https://doi.org/10.1101/178806

Elwood, E., Lim, Z., Naveed, H., & Galea, I. (2017). The effect of systemic inflammation on human brain barrier function. *Brain, Behavior, and Immunity*, *62*, 35–40. https://doi.org/10.1016/j.bbi.2016.10.020

Emanuele, E., Peros, E., Tomaino, C., Feudatari, E., Bernardi, L., Binetti, G., … Geroldi, D. (2004). Apolipoprotein(a) null phenotype is related to a delayed age at onset of Alzheimer's disease. *Neuroscience Letters*, *357*(1), 45–48. https://doi.org/10.1016/j.neulet.2003.12.043

España-Irla, G., Gomes-Osman, J., Cattaneo, G., Albu, S., Cabello-Toscano, M., Solana-Sanchéz, J., … Pascual-Leone, Á. (2021). Associations between cardiorespiratory fitness, cardiovascular risk, and cognition are mediated by structural brain health in midlife. *Journal of the American Heart Association*, *10*(18). https://doi.org/10.1161/JAHA.120.020688

Farooqui, A. A., Liss, L., & Horrocks, L. A. (1988). Neurochemical aspects of Alzheimer's disease: Involvement of membrane phospholipids. *Metabolic Brain Disease*, *3*(1), 19–35. https://doi.org/10.1007/BF01001351

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., … Van Duijn, C. M. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: A meta-analysis. *Journal of the American Medical Association*, *278*(16), 1349–1356. https://doi.org/10.1001/jama.278.16.1349

Fawns-Ritchie, C., & Deary, I. J. (2020). Reliability and validity of the UK Biobank cognitive tests. *PLoS ONE*, *15*(4), 1–24. https://doi.org/10.1371/journal.pone.0231627

Ferguson, A. C., Tank, R., Lyall, L. M., Ward, J., Celis-Morales, C., Strawbridge, R., … Lyall, D. M. (2020). Alzheimer's Disease Susceptibility Gene Apolipoprotein E (APOE) and Blood Biomarkers in UK Biobank (N = 395,769). *Journal of Alzheimer's Disease*, *76*(4), 1541–1551. https://doi.org/10.3233/jad-200338

Ferguson, S. C., Blane, A., Perros, P., McCrimmon, R. J., Best, J. J. K., Wardlaw, J., … Frier, B. M. (2003). Cognitive Ability and Brain Structure in Type 1 Diabetes. *Diabetes*, *52*(1), 149–156. https://doi.org/10.2337/DIABETES.52.1.149

Fernandes, A., Tábuas-Pereira, M., Duro, D., Lima, M., Gens, H., Santiago, B., … Santana, I. (2020). C-reactive protein as a predictor of mild cognitive impairment conversion into Alzheimer's disease dementia. *Experimental Gerontology*, *138*(June), 111004. https://doi.org/10.1016/j.exger.2020.111004

Ferreira, D., Nordberg, A., & Westman, E. (2020). Biological subtypes of Alzheimer disease: A systematic review and meta-analysis. *Neurology*, *94*(10), 436–448. https://doi.org/10.1212/WNL.0000000000009058

Ferreira, D., Verhagen, C., Hernández-Cabrera, J. A., Cavallin, L., Guo, C. J., Ekman, U., … Westman, E. (2017). Distinct subtypes of Alzheimer's disease based on patterns of brain atrophy: Longitudinal trajectories and clinical applications. *Scientific Reports*, *7*(March), 1–13. https://doi.org/10.1038/srep46263

Ferreira, J. P., Kearney Schwartz, A., Zohra, L., Felblinger, J., Boivin, J. M., Bracard, S., … Verger, A. (2017). Memory Alterations and White Matter Hyperintensities in Elderly Patients With Hypertension: The ADELAHYDE-2 Study. *Journal of the American Medical Directors Association*, *18*(5), 451.e13-451.e25. https://doi.org/10.1016/j.jamda.2017.01.008

Ferretti, M. T., Iulita, M. F., Cavedo, E., Chiesa, P. A., Dimech, A. S., Chadha, A. S., … Hampel, H. (2018). Sex differences in Alzheimer disease — The gateway to precision medicine. *Nature Reviews Neurology*, *14*(8), 457–469. https://doi.org/10.1038/s41582-018-0032-9

Filippi, M., Basaia, S., Canu, E., Imperiale, F., Magnani, G., Falautano, M., … Agosta, F. (2020). Changes in functional and structural brain connectome along the Alzheimer's disease continuum. *Molecular Psychiatry*, *25*(1), 230–239. https://doi.org/10.1038/s41380-018-0067-8

Flanagin, A., Frey, T., & Christiansen, S. L. (2021). Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA - Journal of the American Medical Association*, *326*(7), 621–627. https://doi.org/10.1001/jama.2021.13304

Friedman, J. I., Tang, C. Y., De Haas, H. J., Changchien, L., Goliasch, G., Dabas, P., … Narula, J. (2014). Brain imaging changes associated with risk factors for cardiovascular and cerebrovascular disease in asymptomatic patients. *JACC: Cardiovascular Imaging*, *7*(10), 1039–1053. https://doi.org/10.1016/j.jcmg.2014.06.014

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., … Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9), 1026–1034. https://doi.org/10.1093/aje/kwx246

Gadd, D. A., Stevenson, A. J., Hillary, R. F., McCartney, D. L., Wrobel, N., McCafferty, S., … Marioni, R. E. (2021). Epigenetic predictors of lifestyle traits applied to the blood and brain. *Brain Communications*, *3*(2), 1–11. https://doi.org/10.1093/braincomms/fcab082

Galinsky, K. J., Loh, P. R., Mallick, S., Patterson, N. J., & Price, A. L. (2016). Population Structure of UK Biobank and Ancient Eurasians Reveals Adaptation at Genes Influencing Blood Pressure. *American Journal of Human Genetics*, *99*(5), 1130–1139. https://doi.org/10.1016/j.ajhg.2016.09.014

Garcia-Ptacek, S., Eriksdotter, M., Jelic, V., Porta-Etessam, J., Kåreholt, I., & Manzano Palomo, S. (2016). Subjective cognitive impairment: Towards early identification of Alzheimer disease. *Neurologia*, *31*(8), 562–571. https://doi.org/http://dx.doi.org/10.1016/j.nrleng.2013.02.011

Gauthier, S., Kaur, G., Mi, W., Tizon, B., & Levy, E. (2011). Protective mechanisms by cystatin C in neurodegenerative diseases. *Frontiers in Bioscience - Scholar*, *3 S*(2), 541–554. https://doi.org/10.2741/s170

Gennatas, E. D., Avants, B. B., Wolf, D. H., Satterthwaite, T. D., Ruparel, K., Ciric, R., … Gur, R. C. (2017). Age-related effects and sex differences in gray matter density, volume, mass, and cortical thickness from childhood to young adulthood. *Journal of Neuroscience*, *37*(20), 5065–5073. https://doi.org/10.1523/JNEUROSCI.3550-16.2017

Gerdts, E., & Regitz-Zagrosek, V. (2019). Sex differences in cardiometabolic disorders. *Nature Medicine*, *25*(11), 1657–1666. https://doi.org/10.1038/s41591-019-0643-8

Gill, J. M. R. (2020). Linking volume and intensity of physical activity to mortality. *Nature Medicine*, *26*(9), 1332–1334. https://doi.org/10.1038/s41591-020-1019-9

Gong, J., Harris, K., Peters, S. A. E., & Woodward, M. (2021). Sex differences in the association between major cardiovascular risk factors in midlife and dementia: a cohort study using data from the UK Biobank. *BMC Medicine*, *19*(1), 1–11. https://doi.org/10.1186/s12916-021-01980-z

Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., … Thal, L. J. (2004). Mild cognitive impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Arch Neurol*, *61*(1), 59–66. https://doi.org/10.1001/archneur.61.1.59

Guerreiro, M. J. S., & Van Gerven, P. W. M. (2011). Now you see it, now you don't: Evidence for age-dependent and age-independent cross-modal distraction. *Psychology and Aging*, *26*(2), 415–426. https://doi.org/10.1037/a0021507

Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., & Alexopoulos, G. S. (2009). Aging of cerebral white matter: A review of MRI findings. *International*

Journal of Geriatric Psychiatry, *24*(2), 109–117. https://doi.org/10.1002/gps.2087

Gupta, N., Simpkins, A. N., Hitomi, E., Dias, C., & Leigh, R. (2018). White Matter Hyperintensity-Associated Blood-Brain Barrier Disruption and Vascular Risk Factors. *Journal of Stroke and Cerebrovascular Diseases*, *27*(2), 466–471. https://doi.org/10.1016/j.jstrokecerebrovasdis.2017.09.026

Hachiya, T., Kamatani, Y., Takahashi, A., Hata, J., Furukawa, R., Shiwa, Y., … Shimizu, A. (2017). Genetic Predisposition to Ischemic Stroke: A Polygenic Risk Score. *Stroke*, *48*(2), 253–258. https://doi.org/10.1161/STROKEAHA.116.014506

Hagenaars, S. P., Cox, S. R., Hill, W. D., Davies, G., Liewald, D. C. M., Harris, S. E., … Deary, I. J. (2018). Genetic contributions to Trail Making Test performance in UK Biobank. *Molecular Psychiatry*, *23*(7), 1575–1583. https://doi.org/10.1038/mp.2017.189

Hagenaars, S. P., Harris, S. E., Davies, G., Hill, W. D., Liewald, D. C. M., Ritchie, S. J., … Deary, I. J. (2016). Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. *Molecular Psychiatry*, *21*(11), 1624–1632. https://doi.org/10.1038/mp.2015.225

Halliday, M. R., Rege, S. V., Ma, Q., Zhao, Z., Miller, C. A., Winkler, E. A., & Zlokovic, B. V. (2016). Accelerated pericyte degeneration and blood-brain barrier breakdown in apolipoprotein E4 carriers with Alzheimer's disease. *Journal of Cerebral Blood Flow and Metabolism*, *36*(1), 216–227. https://doi.org/10.1038/jcbfm.2015.44

Hampel, H., Schneider, L. S., Giacobini, E., Kivipelto, M., Sindi, S., Dubois, B., … Lista, S. (2015). Advances in the therapy of Alzheimer's disease: targeting amyloid beta and tau and perspectives for the future. *Expert Review of Neurotherapeutics*, *15*(1), 83–105. https://doi.org/10.1586/14737175.2015.995637

Harada, C. N., Natelson Love, M. C., & Triebel, K. (2013). Normal Cognitive Aging. *Clinics in Geriatric Medicine*, *29*(4), 737–752. https://doi.org/10.1016/j.cger.2013.07.002.Normal

Hedden, T., & Gabrieli, J. De. (2005). Healthy and pathological processes in adult development: new evidence from neuroimaging of the aging brain. *Current Opinion in Neurology*, *18*(6), 740–747. https://doi.org/10.1097/01.wco.0000189875.29852.48

Hegele, R. A., & Dichgans, M. (2010). Advances in stroke 2009: Update on the genetics of stroke and cerebrovascular disease 2009. *Stroke*, *41*(2). https://doi.org/10.1161/STROKEAHA.109.571034

Henriksen, K., Bryant, S. E. O., Hampel, H., Trojanowski, J. Q., Montine, T. J., Wyss-coray, T., … Interest, B. B. (2014). The future of blood-based biomarkers for Alzheimer's disease. *Alzheimer's & Dementia*, *10*, 115–131.

Henson, R. N., Suri, S., Knights, E., Rowe, J. B., Kievit, R. A., Lyall, D. M., … Fisher, S. E. (2020). Effect of apolipoprotein E polymorphism on cognition and brain in the Cambridge Centre for Ageing and Neuroscience cohort. *Brain and Neuroscience Advances*, *4*, 239821282096170. https://doi.org/10.1177/2398212820961704

Higgins, V., Nazroo, J., & Brown, M. (2019). Pathways to ethnic differences in obesity: The role of migration, culture and socio-economic position in the UK. *SSM - Population Health*, *7*(March), 100394. https://doi.org/10.1016/j.ssmph.2019.100394

Hill-Briggs, F., Adler, N. E., Berkowitz, S. A., Chin, M. H., Gary-Webb, T. L., Navas-Acien, A., … Haire-Joshu, D. (2021). Social determinants of health and diabetes: A scientific review. *Diabetes Care*, *44*(1), 258–279. https://doi.org/10.2337/dci20-0053

Hollands, C., Bartolotti, N., & Lazarov, O. (2016). Alzheimer's Disease and Hippocampal Adult Neurogenesis; Exploring Shared Mechanisms. *Frontiers in Neuroscience*, *10*(May), 1–8. https://doi.org/10.3389/fnins.2016.00178

Hooghiemstra, A. M., Bertens, A. S., Leeuwis, A. E., Bron, E. E., Bots, M. L., Brunner-La Rocca, H. P., … Van Der Flier, W. M. (2017). The Missing Link in the Pathophysiology of Vascular Cognitive Impairment: Design of the Heart-Brain Study. *Cerebrovascular Diseases Extra*, *7*(3), 140–152. https://doi.org/10.1159/000480738

Hou, Y., Dan, X., Babbar, M., Wei, Y., Hasselbalch, S. G., Croteau, D. L., & Bohr, V. A. (2019). Ageing as a risk factor for neurodegenerative disease. *Nature Reviews Neurology*, *15*(10), 565–581. https://doi.org/10.1038/s41582-019-0244-7

Howard, G., Wagenknecht, L. E., Burke, G. L., Diez-Roux, A., Evans, G. W., McGovern, P., … Tell, G. S. (1998). Cigarette smoking and progression of atherosclerosis: The atherosclerosis risk in communities (ARIC) study. *Journal of the American Medical Association*, *279*(2), 119–124. https://doi.org/10.1001/jama.279.2.119

Hu, L., Liu, B., & Li, Y. (2020). Ranking sociodemographic, health behavior, prevention, and environmental factors in predicting neighborhood cardiovascular health: A Bayesian machine learning approach. *Preventive Medicine*, *141*(July), 1–4. https://doi.org/10.1016/j.ypmed.2020.106240

Huo, Z., Yu, L., Yang, J., Zhu, Y., Bennett, D. A., & Zhao, J. (2020). Brain and blood metabolome for Alzheimer's dementia: findings from a targeted metabolomics analysis. *Neurobiology of Aging*, *86*, 123–133. https://doi.org/10.1016/j.neurobiolaging.2019.10.014

Huynh, R. A., & Mohan, C. (2017). Alzheimer's disease: Biomarkers in the genome, blood, and cerebrospinal fluid. *Frontiers in Neurology*, *8*(MAR). https://doi.org/10.3389/fneur.2017.00102

Iturria-Medina, Y., Hachinski, V., & Evans, A. C. (2017). The vascular facet of late-onset Alzheimer's disease: An essential factor in a complex multifactorial disorder. *Current Opinion in Neurology*, *30*(6), 623–629. https://doi.org/10.1097/WCO.0000000000000497

Iwamoto, T., Watanabe, D., Umahara, T., Sakurai, H., Hanyu, H., & Kanaya, K. (2004). Dual inverse effects of lipoprotein(a) on the dementia process in Japanese late-onset Alzheimer's disease. In *PSYCHOGERIATRICS* (Vol. 4).

Jacobs, H. I. L., Clerx, L., Gronenschild, E. H. B. M., Aalten, P., & Verhey, F. R. J. (2014). White matter hyperintensities are positively associated with cortical thickness in Alzheimer's disease. *Journal of Alzheimer's Disease*, *39*(2), 409–422. https://doi.org/10.3233/JAD-131232

Jacobs, H. I. L., Van Boxtel, M. P. J., Jolles, J., Verhey, F. R. J., & Uylings, H. B. M. (2012). Parietal cortex matters in Alzheimer's disease: An overview of structural, functional and metabolic findings. *Neuroscience and Biobehavioral Reviews*, *36*(1), 297–309. https://doi.org/10.1016/j.neubiorev.2011.06.009

Jacobs, H. I. L., Visser, P. J., Van Boxtel, M. P. J., Frisoni, G. B., Tsolaki, M., Papapostolou, P., … Verhey, F. R. J. (2012). The association between white matter hyperintensities and executive decline in mild cognitive impairment is network dependent. *Neurobiology of Aging*, *33*(1), 201.e1-201.e8. https://doi.org/10.1016/j.neurobiolaging.2010.07.015

Kamstrup, P. R., Tybjærg-hansen, A., Steffensen, R., & Nordestgaard, B. G. (2009). Genetically Elevated Lipoprotein (a). *The Journal of the American Medical Association*, *301*(22), 2331–2339.

Kaufmann, T., & Meer, D. van der. (2020). Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat Neurosci*, *22*(10), 1617–1623. https://doi.org/10.1038/s41593-019-0471-7.Common

Kaur, G., & Levy, E. (2012). Cystatin C in Alzheimer's disease. *Frontiers in Molecular Neuroscience*, *5*(JULY 2012), 1–11. https://doi.org/10.3389/fnmol.2012.00079

Khera, A. V., Chaffin, M., & Aragam, K. G. (2018). Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet.*, *176*(1), 100–106. https://doi.org/10.1038/s41588-018-0183-z.Genome-wide

Khera, A., Vega, G. L., Das, S. R., Ayers, C., McGuire, D. K., Grundy, S. M., & De Lemos, J. A. (2009). Sex differences in the relationship between c-reactive protein and body fat. *Journal of Clinical Endocrinology and Metabolism*, *94*(9), 3251–3258. https://doi.org/10.1210/jc.2008-2406

Khoury, R., & Ghossoub, E. (2019). Diagnostic biomarkers of Alzheimer's disease: A state-of-the-art review. *Biomarkers in Neuropsychiatry*, *1*(November), 100005. https://doi.org/10.1016/j.bionps.2019.100005

Khramtsova, E. A., Davis, L. K., & Stranger, B. E. (2019). The role of sex in the genomics of human complex traits. *Nature Reviews Genetics*, *20*(3), 173–190. https://doi.org/10.1038/s41576-018-0083-1

Kim, J., Basak, J. M., & Holtzman, D. M. (2009). The Role of Apolipoprotein E in Alzheimer's Disease. *Neuron*, *63*(3), 287–303. https://doi.org/10.1016/j.neuron.2009.06.026

King, K. S., Peshock, R. M., Rossetti, H. C., McColl, R. W., Ayers, C. R., Hulsey, K. M., & Das, S. R. (2014). Effect of normal aging versus hypertension, abnormal body mass index, and diabetes mellitus on white matter hyperintensity volume. *Stroke*, *45*(1), 255–257. https://doi.org/10.1161/Strokeaha.113.003602

Kivipelto, M., Solomon, A., Ahtiluoto, S., Ngandu, T., Lehtisalo, J., Antikainen, R., … Soininen, H. (2013). The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER): Study design and progress. *Alzheimer's and Dementia*, *9*(6), 657–665. https://doi.org/10.1016/j.jalz.2012.09.012

Knopman, D. S., Penman, A. D., Catellier, D. J., Coker, L. H., Shibata, D. K., Sharrett, A. R., & Mosley, T. H. (2011). Vascular risk factors and longitudinal changes on brain MRI: The ARIC study. *Neurology*, *76*(22), 1879–1885. https://doi.org/10.1212/WNL.0b013e31821d753f

Knopman, D. S., Mosley, T. H., Catellier, D. J., & Sharrett, A. R. (2005). Cardiovascular risk factors and cerebral atrophy in a middle-aged cohort.

Neurology, *65*(6), 876–881. https://doi.org/10.1212/01.wnl.0000176074.09733.a8

Kolbeinsson, A., Filippi, S., Panagakis, Y., Matthews, P. M., Elliott, P., Dehghan, A., & Tzoulaki, I. (2020). Accelerated MRI-predicted brain ageing and its associations with cardiometabolic and brain disorders. *Scientific Reports*, *10*(1), 1–9. https://doi.org/10.1038/s41598-020-76518-z

Krieger, N. (2002). A glossary for social epidemiology. *Epidemiological Bulletin*, *23*(1), 7–11. https://doi.org/10.1046/j.1365-2214.2002.0255g.x

Krogager, M. L., Skals, R. K., Appel, E. V. R., Schnurr, T. M., Engelbrechtsen, L., Have, C. T., … Weeke, P. E. (2019). Hypertension genetic risk score is associated with burden of coronary heart disease among patients referred for coronary angiography. *PLoS ONE*, *13*(12), 1–17. https://doi.org/10.1371/journal.pone.0208645

Kronenberg, F. (2016). Human Genetics and the Causal Role of Lipoprotein(a) for Various Diseases. *Cardiovascular Drugs and Therapy*, *30*(1), 87–100. https://doi.org/10.1007/s10557-016-6648-3

Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., … Pericak-Vance, M. A. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*, *51*(3), 414–430. https://doi.org/10.1038/s41588-019-0358-2

Kunutsor, S. K., Khan, H., Nyyssönen, K., & Laukkanen, J. A. (2016). Is lipoprotein (a) protective of dementia? *European Journal of Epidemiology*, *31*(11), 1149–1152. https://doi.org/10.1007/s10654-016-0184-0

Laffoon, S. B., Doecke, J. D., Roberts, A. M., Vance, J. A., Reeves, B. D., Pertile, K. K., … Roberts, B. R. (2020). Discovery of plasma protein biomarkers related to Alzheimer's disease, sex and APOE genotype. *MedRxiv*, (404). https://doi.org/10.1101/2020.04.15.20066647

Lamar, M., Boots, E. A., Arfanakis, K., Barnes, L. L., & Schneider, J. A. (2020). Common Brain Structural Alterations Associated with Cardiovascular Disease Risk Factors and Alzheimer's Dementia: Future Directions and Implications. *Neuropsychology Review*, *30*(4), 546–557. https://doi.org/10.1007/s11065-020-09460-6

Lastra, G., Syed, S., Kurukulasuriya, L. R., Manrique, C., & Sowers, J. R. (2014). Type 2 diabetes mellitus and hypertension: An update. *Endocrinology and Metabolism Clinics of North America*, *43*(1), 103–122. https://doi.org/10.1016/j.ecl.2013.09.005

Lau, E. S., Paniagua, S. M., Guseh, J. S., Bhambhani, V., Zanni, M. V., Courchesne, P., … Ho, J. E. (2019). Sex Differences in Circulating Biomarkers of Cardiovascular Disease. *Journal of the American College of Cardiology*, *74*(12), 1543–1553. https://doi.org/10.1016/j.jacc.2019.06.077

Lau, W. L., Fisher, M., Greenia, D., Floriolli, D., Fletcher, E., Singh, B., … Paganini-Hill, A. (2020). Cystatin C, cognition, and brain MRI findings in 90+-year-olds. *Neurobiology of Aging*, *93*, 78–84. https://doi.org/10.1016/j.neurobiolaging.2020.04.022

Laurin, D., Curb, J. D., Masaki, K. H., White, L. R., & Launer, L. J. (2009). Midlife C-reactive protein and risk of cognitive decline: A 31-year follow-up. *Neurobiol Aging*. https://doi.org/10.1016/j.neurobiolaging.2008.01.008.Midlife

Leeson, G. W. (2018). The Growth, Ageing and Urbanisation of our World. *Journal of Population Ageing*, *11*(2), 107–115. https://doi.org/10.1007/s12062-018-9225-7

Leonenko, G., Sims, R., Shoai, M., Frizzati, A., Bossù, P., Spalletta, G., … Escott-Price, V. (2019). Polygenic risk and hazard scores for Alzheimer's disease prediction. *Annals of Clinical and Translational Neurology*, *6*(3), 456–465. https://doi.org/10.1002/acn3.716

Levine, M. E. (2013). Modeling the rate of senescence: Can estimated biological age predict mortality more accurately than chronological age? *Journals of*

Gerontology - Series A Biological Sciences and Medical Sciences, *68*(6), 667–674. https://doi.org/10.1093/gerona/gls233

Levine, Lu, Quach, & Chen. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *AGING*, (4), 573–591. https://doi.org/10.1145/2820783.2820789

Lew, J., Sanghavi, M., Ayers, C. R., McGuire, D. K., Omland, T., Atzler, D., … De Lemos, J. A. (2017). Sex-Based Differences in Cardiometabolic Biomarkers. *Circulation*, *135*(6), 544–555. https://doi.org/10.1161/CIRCULATIONAHA.116.023005

Li, J., Wang, Y. J., Zhang, M., Xu, Z. Q., Gao, C. Y., Fang, C. Q., … Zhou, H. D. (2011). Vascular risk factors promote conversion from mild cognitive impairment to Alzheimer disease. *Neurology*, *76*(17), 1485–1491. https://doi.org/10.1212/WNL.0b013e318217e7a4

Li, X., Meng, X., He, Y., Spiliopoulou, A., Timofeeva, M., Wei, W. Q., … Theodoratou, E. (2019). Genetically determined serum urate levels and cardiovascular and other diseases in UK Biobank cohort: A phenome-wide mendelian randomization study. *PLoS Medicine*, *16*(10), 1–20. https://doi.org/10.1371/journal.pmed.1002937

Lippi, G., Danese, E., & Favaloro, E. J. (2019). Vascular Disease and Dementia: Lipoprotein (a) as a Neglected Link. *Seminars in Thrombosis and Hemostasis*, *45*(5), 544–547. https://doi.org/10.1055/s-0039-1687907

Logue, M. W., Panizzon, M. S., Elman, J. A., Gillespie, N. A., Hatton, S. N., Gustavson, D. E., … Kremen, W. S. (2019). Use of an Alzheimer's disease polygenic risk score to identify mild cognitive impairment in adults in their 50s. *Molecular Psychiatry*, *24*(3), 421–430. https://doi.org/10.1038/s41380-018-0030-8

Lorenzo, C., OKOLOISE, M., Williams, K., Stern, M., & Haffner, S. (2003). The Metabolic Syndrome as Predictor of Type 2 Diabetes. *Diabetes Care*, *26*(11).

Lupton, M. K., Strike, L., Hansell, N. K., Wen, W., Mather, K. A., Armstrong, N. J., … Wright, M. J. (2016). The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume the Alzheimer's Disease Neuroimaging Initiative HHS Public Access. *Neurobiol Aging*, *40*, 68–77. https://doi.org/10.1016/j.neurobiolaging.2015.12.023

Lyall, D. M., Quinn, T., Lyall, L. M., Ward, J., Anderson, J., Smith, D. J., & Stewart, W. (2021). Quantifying bias in psychological and physical health in the UK Biobank imaging sub-sample. *PsyArXiv*, (July), 1–17.

Lyall, D. M., Celis-Morales, C. A., Anderson, J., Gill, J. M. R., Mackay, D. F., McIntosh, A. M., … Pell, J. P. (2017). Associations between single andmultiple cardiometabolic diseases and cognitive abilities in 474 129 UK Biobank participants. *European Heart Journal*, *38*(8), 577–583. https://doi.org/10.1093/eurheartj/ehw528

Lyall, D. M., Cullen, B., Allerhand, M., Smith, D. J., Mackay, D., Evans, J., … Pell, J. P. (2016). Cognitive test scores in UK biobank: Data reduction in 480,416 participants and longitudinal stability in 20,346 participants. *PLoS ONE*, *11*(4), 1–10. https://doi.org/10.1371/journal.pone.0154222

Lyall, D. M., Ward, J., Ritchie, S. J., Davies, G., Cullen, B., Celis, C., … Pell, J. P. (2016). Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants. *Age and Ageing*, *45*(4), 511–517. https://doi.org/10.1093/ageing/afw068

Lyall, D. M., Royle, N. A., Harris, S. E., Bastin, M. E., & Muñoz-Maniega, S. (2013). Alzheimer's Disease Susceptibility Genes APOE and TOMM40, and Hippocampal Volumes in the Lothian Birth Cohort 1936. *PLOS One*, Vol. 388, pp. 1–14.

Lyall, L. M., Cullen, B., Lyall, D. M., Leighton, S. P., Siebert, S., Smith, D. J., & Cavanagh, J. (2019). The associations between self-reported depression, self-reported chronic inflammatory conditions and cognitive abilities in UK Biobank. *European Psychiatry*, *60*, 63–70. https://doi.org/10.1016/j.eurpsy.2019.05.007

Mack, S., Coassin, S., Rueedi, R., Yousri, N. A., Seppälä, I., Gieger, C., … Kronenberg, F. (2017). A genome-wide association meta-analysis on lipoprotein (a) concentrations adjusted for apolipoprotein (a) isoforms. *Journal of Lipid Research*, *58*(9), 1834–1844. https://doi.org/10.1194/jlr.M076232

Mannucci, E., Dicembrini, I., Lauria, A., & Pozzilli, P. (2013). Is glucose control important for prevention of cardiovascular disease in diabetes? *Diabetes Care*, *36*(SUPPL.2). https://doi.org/10.2337/dcS13-2018

Marchant, N. L., Reed, B. R., DeCarli, C. S., Madison, C. M., Weiner, M. W., Chui, H. C., & Jagust, W. J. (2012). Cerebrovascular disease, beta-amyloid, and cognition in aging. *Neurobiology of Aging*, *33*(5), 1006.e25-1006.e36. https://doi.org/10.1016/j.neurobiolaging.2011.10.001

Marioni, R. E., Campbell, A., Hagenaars, S. P., Nagy, R., Amador, C., Hayward, C., … Deary, I. J. (2017). Genetic Stratification to Identify Risk Groups for Alzheimer's Disease. *Journal of Alzheimer's Disease*, *57*(1), 275–283. https://doi.org/10.3233/JAD-161070

Marioni, R. E., Harris, S. E., Zhang, Q., McRae, A. F., Hagenaars, S. P., Hill, W. D., … Visscher, P. M. (2018). GWAS on family history of Alzheimer's disease. *Translational Psychiatry*, *8*(1), 0–6. https://doi.org/10.1038/s41398-018-0150-6

Mathews, & Levy. (2016). Cystatin C in aging and in Alzheimer's disease. *Ageing Res Rev.*, *176*(3), 139–148. https://doi.org/10.1016/j.arr.2016.06.003.Cystatin

McKhann, G., Knopman, D. S., Chertkow, H., Hymann, B., Jack, C. R., Kawas, C., … Phelphs, C. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dementia*, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005.The

Melzer, D., Pilling, L. C., & Ferrucci, L. (2020). The genetics of human ageing. *Nature Reviews Genetics*, *21*(2), 88–101. https://doi.org/10.1038/s41576-019-0183-6

Michael Marmot. (2010). Fair society, healthy lives: the Marmot review; strategic review of health inequalities in England post-2010. *The Marmot Review*. https://doi.org/10.1136/bmj.c1191

Mohanty, R., Mårtensson, G., Poulakis, K., Muehlboeck, J. S., Rodriguez-Vieitez, E., Chiotis, K., … Westman, E. (2020). Towards harmonizing subtyping methods for neuroimaging studies in Alzheimer's disease. *MedRxiv*. https://doi.org/10.1101/2020.04.19.20064881

Mooser, V., Helbecque, N., Miklossy, J., Santica, ;, Marcovina, M., Nicod, P., & Amouyel, P. (2000). *Interactions between Apolipoprotein E and Apolipoprotein(a) in Patients with Late-Onset Alzheimer Disease*.

Morgan, K. N., Derby, C. A., & Gleason, C. E. (2018). Cognitive Changes with Reproductive Aging, Perimenopause, and Menopause. *Obstetrics and Gynecology Clinics of North America*, *45*(4), 751–763. https://doi.org/10.1016/j.ogc.2018.07.011

Mormino, E. C., Papp, K. V., Rentz, D. M., Schultz, A. P., LaPoint, M., Amariglio, R., … Sperling, R. A. (2016). Heterogeneity in Suspected Non-Alzheimer Disease Pathophysiology Among Clinically Normal Older Individuals. *JAMA Neurology*, *73*(10), 1185. https://doi.org/10.1001/jamaneurol.2016.2237

Mormino, E. C., & Sperling, R. A. (2016). Polygenic risk of Alzheimer disease is associated with early- and late-life processes. *Neurology*, *87*(5), 481–488. https://doi.org/10.1212/WNL.0000000000002922

Mosconi, L., Berti, V., Quinn, C., McHugh, P., Petrongolo, G., Varsavsky, I., … Brinton, R. D. (2017). Sex differences in Alzheimer risk. *Neurology*, *89*(13), 1382–1390. https://doi.org/10.1212/WNL.0000000000004425

Mosley, T. H., Knopman, D. S., Catellier, D. J., Bryan, N., Hutchinson, R. G., Grothues, C. A., … Szklo, M. (2005). Cerebral MRI findings and cognitive functioning: The Atherosclerosis Risk in Communities Study. *Neurology*, *64*(12), 2056–2062. https://doi.org/10.1212/01.WNL.0000165985.97397.88

Mrdjen, D., Fox, E. J., Bukhari, S. A., Montine, K. S., Bendall, S. C., & Montine, T. J. (2019). The basis of cellular and regional vulnerability in Alzheimer's disease. *Acta Neuropathologica*, *138*(5), 729–749. https://doi.org/10.1007/s00401-019-02054-4

Munafò, M. R. (2006). Candidate gene studies in the 21st century: meta-analysis, mediation, moderation. *Genes, Brain, and Behavior*, *5 Suppl 1*, 3–8. https://doi.org/10.1111/j.1601-183X.2006.00188.x

Munafò, M. R., & Davey Smith, G. (2018). Repeating experiments is not enough. *Nature*, *553*(7689), 399–401. https://doi.org/10.1038/d41586-018-01023-3

Murray, A. D., Staff, R. T., Shenkin, S. D., Deary, I. J., Starr, J. M., & Whalley, L. J. (2005). Brain white matter hyperintensities: Relative importance of vascular risk factors in nondemented elderly people. *Radiology*, *237*(1), 251–257. https://doi.org/10.1148/radiol.2371041496

Muse, E. D., Wineinger, N. E., & Spencer, E. G. (2018). Validation of a genetic risk score for atrial fibrillation: A prospective multicenter cohort study. *PLOS Medicine*, 1–12. Retrieved from http://repositorio.ug.edu.ec/handle/redug/8495

Nations, U. (2019). World population prospects 2019. In *Department of Economic and Social Affairs. World Population Prospects 2019*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12283219

O'Bryant, S. E., Mielke, M. M., Rissman, R. A., Lista, S., Vanderstichele, H., Zetterberg, H., … Hampel, H. (2017). Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic. *Alzheimer's and Dementia*, *13*(1), 45–58. https://doi.org/10.1016/j.jalz.2016.09.014

O'Bryant, S. E., Mielke, M. M., Rissman, R. A., Lista, S., Vanderstichele, H., Zetterberg, H., … Hampel, H. (2017). Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic. *Alzheimer's and Dementia*, *13*(1), 45–58. https://doi.org/10.1016/j.jalz.2016.09.014

Olazarán, J., Gil-De-Gómez, L., Rodríguez-Martín, A., Valentí-Soler, M., Frades-Payo, B., Marín-Muñoz, J., … Calero, M. (2015). A blood-based, 7-metabolite signature for the early diagnosis of Alzheimer's disease. *Journal of Alzheimer's Disease*, *45*(4), 1157–1173. https://doi.org/10.3233/JAD-142925

Pan, Y., Li, H., Wang, Y., Meng, X., & Wang, Y. (2019). Causal effect of Lp(a) [lipoprotein(a)] level on ischemic stroke and alzheimer disease a mendelian randomization study. *Stroke*, *50*(12), 3532–3539. https://doi.org/10.1161/STROKEAHA.119.026872

Penninx, B. W. J. H. (2017). Depression and cardiovascular disease: Epidemiological evidence on their linking mechanisms. *Neuroscience and Biobehavioral Reviews*, *74*, 277–286. https://doi.org/10.1016/j.neubiorev.2016.07.003

Petrea, R. E., O'Donnell, A., Beiser, A. S., Habes, M., Aparicio, H., Decarli, C., … Romero, J. R. (2020). Mid to Late Life Hypertension Trends and Cerebral Small Vessel Disease in the Framingham Heart Study. *Hypertension*, 707–714. https://doi.org/10.1161/HYPERTENSIONAHA.120.15073

Petrie, J. R., Guzik, T. J., & Touyz, R. M. (2018). Diabetes, Hypertension, and Cardiovascular Disease: Clinical Insights and Vascular Mechanisms. *Canadian Journal of Cardiology*, *34*(5), 575–584. https://doi.org/10.1016/j.cjca.2017.12.005

Pham, T. M., Petersen, I., Walters, K., Raine, R., Manthorpe, J., Mukadam, N., & Cooper, C. (2018). Trends in dementia diagnosis rates in uk ethnic groups: Analysis of uk primary care data. *Clinical Epidemiology*, *10*, 949–960. https://doi.org/10.2147/CLEP.S152647

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Profaci, C. P., Munji, R. N., Pulido, R. S., & Daneman, R. (2020). The blood-brain barrier in health and disease: Important unanswered questions. *Journal of Experimental Medicine*, *217*(4), 1–16. https://doi.org/10.1084/jem.20190062

Proitsi, P., Kim, M., Whiley, L., Simmons, A., Sattlecker, M., Velayudhan, L., … Legido-Quigley, C. (2017). Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. *Alzheimer's and Dementia*, *13*(2), 140–151. https://doi.org/10.1016/j.jalz.2016.08.003

Proitsi, P., Kim, M., Whiley, L., Simmons, A., Sattlecker, M., Velayudhan, L., … Legido-Quigley, C. (2017). Association of blood lipids with Alzheimer's disease: A comprehensive lipidomics analysis. *Alzheimer's and Dementia*, *13*(2), 140–151. https://doi.org/10.1016/j.jalz.2016.08.003

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rabeya, K., Neela, M. M., Ahsan, R., Biswas, B. K., & Hossain, M. M. (2021). A Comparative Study of Memory, Attention, Cognition and Oestradiol in Pre and Postmenopausal Women in Bangladesh. *Bangladesh Pharmaceutical Journal*, *24*(1), 26–32. https://doi.org/10.3329/bpj.v24i1.51632

Radanovic, M. (2020). Cognitive reserve: an evolving concept. *International Psychogeriatrics*, *32*(1), 7–9. https://doi.org/10.1017/S1041610219001947

Raghavan, N. S., Vardarajan, B., & Mayeux, R. (2019). Genomic variation in educational attainment modifies Alzheimer disease risk. *Neurology: Genetics*, *5*(2). https://doi.org/10.1212/NXG.0000000000000310

Reitz, C. (2013). Dyslipidemia and the risk of alzheimer's disease. *Current Atherosclerosis Reports*, *15*(3). https://doi.org/10.1007/s11883-012-0307-3

Riveros-Mckay, F., Weale, M. E., Moore, R., Selzam, S., Krapohl, E., Sivley, R. M., … Donnelly, P. (2021). Integrated Polygenic Tool Substantially Enhances

Coronary Artery Disease Prediction. *Circulation: Genomic and Precision Medicine*, pp. 192–200. https://doi.org/10.1161/CIRCGEN.120.003304

Rojas, S., Brugulat-Serrat, A., Bargalló, N., Minguillón, C., Tucholka, A., Falcon, C., … Gispert, J. D. (2018). Higher prevalence of cerebral white matter hyperintensities in homozygous APOE-ε4 allele carriers aged 45–75: Results from the ALFA study. *Journal of Cerebral Blood Flow and Metabolism*, *38*(2), 250–261. https://doi.org/10.1177/0271678X17707397

Rossor, M., & Knapp, M. (2015). Can we model a cognitive footprint of interventions and policies to help to meet the global challenge of dementia? *The Lancet*, *386*(9997), 1008–1010. https://doi.org/10.1016/S0140-6736(15)60248-3

Ryu, W. S., Woo, S. H., Schellingerhout, D., Chung, M. K., Kim, C. K., Jang, M. U., … Kim, D. E. (2014). Grading and interpretation of white matter hyperintensities using statistical maps. *Stroke*, *45*(12), 3567–3575. https://doi.org/10.1161/STROKEAHA.114.006662

Sachdev, P. S., Thalamuthu, A., Mather, K. A., Ames, D., Wright, M. J., Wen, W., … Lemmon, C. (2016). White Matter Hyperintensities Are under Strong Genetic Influence. *Stroke*, *47*(6), 1422–1428. https://doi.org/10.1161/STROKEAHA.116.012532

Salami, A., Eriksson, J., Nilsson, L. G., & Nyberg, L. (2012). Age-related white matter microstructural differences partly mediate age-related decline in processing speed but not cognition. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1822*(3), 408–415. https://doi.org/10.1016/j.bbadis.2011.09.001

Schilling, S., Tzourio, C., Dufouil, C., Zhu, Y., Berr, C., Alpérovitch, A., … Debette, S. (2014). Plasma lipids and cerebral small vessel disease. *Neurology*, *83*(20), 1844–1852. https://doi.org/10.1212/WNL.0000000000000980

Schmidt, A. F., Swerdlow, D. I., Holmes, M. V., Patel, R. S., Fairhurst-Hunter, Z., Lyall, D. M., … Sattar, N. (2017). PCSK9 genetic variants and risk of type 2

diabetes: a mendelian randomisation study. *The Lancet Diabetes and Endocrinology*, *5*(2), 97–105. https://doi.org/10.1016/S2213-8587(16)30396-5

Schmidt, R., Schmidt, H., Curb, J. D., Masaki, K., White, L. R., & Launer, L. J. (2002). Early inflammation and dementia: A 25-year follow-up of the Honolulu-Asia Aging Study. *Annals of Neurology*, *52*(2), 168–174. https://doi.org/10.1002/ana.10265

Shi, Z., Yu, H., Wu, Y., Ford, M., Perschon, C., Wang, C., … Xu, J. (2019). Genetic risk score modifies the effect of APOE on risk and age onset of Alzheimer's disease. *Clinical Genetics*, *95*(2), 302–309. https://doi.org/10.1111/cge.13479

Slunecka, J. L., van der Zee, M. D., Beck, J. J., Johnson, B. N., Finnicum, C. T., Pool, R., … Ehli, E. A. (2021). Implementation and implications for polygenic risk scores in healthcare. *Human Genomics*, *15*(1), 1–18. https://doi.org/10.1186/s40246-021-00339-y

Smith, G. D., Hart, C., Blane, D., Gillis, C., & Hawthorne, V. (1997). Lifetime socioeconomic position and mortality: prospective observational study. *British Medical Journal*, *314*. https://doi.org/10.1136/bmj.314.7080.553

Solfrizzi, V., Panza, F., Colacicco, A. M., Capurso, C., Basile, A. M., & Capurso, A. (2002). Lipoprotein(a), apolipoprotein E genotype, and risk of Alzheimer's disease. *J Neurol Neurosurg Psychiatry*, 732–736.

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., … Van Loenhoud, A. C. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's and Dementia*, *16*(9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219

Stillman, C. M., Weinstein, A. M., Marsland, A. L., Gianaros, P. J., & Erickson, K. I. (2017). Body-brain connections: The effects of obesity and behavioral interventions on neurocognitive aging. *Frontiers in Aging Neuroscience*, *9*(MAY), 1–18. https://doi.org/10.3389/fnagi.2017.00115

Stocker, H., Möllers, T., Perna, L., & Brenner, H. (2018). The genetic risk of Alzheimer's disease beyond APOE ε4: systematic review of Alzheimer's genetic risk scores. *Translational Psychiatry*, *8*(1). https://doi.org/10.1038/s41398-018-0221-8

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., … Collins, R. (2015). UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLoS Medicine*, *12*(3), 1–10. https://doi.org/10.1371/journal.pmed.1001779

Timothy A. Salthouse. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin*, *23*(1), 753–784. https://doi.org/10.1037/a0023262.Neuroanatomical

Toledo, J. B., Arnold, S. E., Raible, K., Brettschneider, J., Xie, S. X., Grossman, M., … Trojanowski, J. Q. (2013). Contribution of cerebrovascular disease in autopsy confirmed neurodegenerative disease cases in the National Alzheimer's Coordinating Centre. *Brain*, *136*(9), 2697–2706. https://doi.org/10.1093/brain/awt188

Toledo, J. B., Shaw, L. M., & Trojanowski, J. Q. (2013). amyloid beta measurements - a desired but elusive Alzheimer's disease biomarker. *Alzheimer's Research & Therapy 2013*.

Tosto, G., Zimmerman, M. E., Carmichael, O. T., & Brickman, A. M. (2014). Predicting Aggressive Decline in Mild Cognitive Impairment. *JAMA Neurology*, *71*(7), 872. https://doi.org/10.1001/jamaneurol.2014.667

Tosto, G., Zimmerman, M. E., Hamilton, J. L., Carmichael, O. T., & Brickman, A. M. (2015). The effect of white matter hyperintensities on neurodegeneration in mild cognitive impairment. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, *11*, 1510–1519. https://doi.org/10.1016/j.jalz.2015.05.014

Trampush, J. W., Yang, M. L. Z., Yu, J., Knowles, E., Davies, G., Liewald, D. C., … Lencz, T. (2017). GWAS meta-analysis reveals novel loci and genetic correlates

for general cognitive function: A report from the COGENT consortium. *Molecular* Psychiatry, 22(3), 336–345. https://doi.org/10.1038/mp.2016.244

Trinder, M., Uddin, M. M., Finneran, P., Aragam, K. G., & Natarajan, P. (2021). Clinical Utility of Lipoprotein(a) and *LPA* Genetic Risk Score in Risk Prediction of Incident Atherosclerotic Cardiovascular Disease. JAMA Cardiology, 6(3), 287–295. https://doi.org/10.1001/jamacardio.2020.5398

Tušek-Bunc, K., & Petek, D. (2016). Comorbidities and characteristics of coronary heart disease patients: Their impact on health-related quality of life. *Health and Quality of Life Outcomes*, *14*(1), 1–12. https://doi.org/10.1186/s12955-016-0560-1

Tyrrell, J., Zheng, J., Beaumont, R., Hinton, K., Richardson, T. G., Wood, A. R., … Tilling, K. (2021). Genetic predictors of participation in optional components of UK Biobank. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-21073-y

van de Haar, H., & Burgmans, S. (2017). Blood-brain barrier leakage in early Alzheimer disease. *Radiology*, *282*(3), 923–924. https://doi.org/10.1148/radiol.2017162578

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2006). The Concept Shifting Test: adult normative data. *Psychological Assessment*, *18*(4), 424–432. https://doi.org/10.1037/1040-3590.18.4.424

Van Dijk, E. J., Breteler, M. M. B., Schmidt, R., Berger, K., Nilsson, L. G., Oudkerk, M., … Hofman, A. (2004). The association between blood pressure, hypertension, and cerebral white matter lesions: Cardiovascular determinants of dementia study. *Hypertension*, *44*(5), 625–630. https://doi.org/10.1161/01.HYP.0000145857.98904.20

Varma, V. R., Oommen, A. M., Varma, S., Casanova, R., An, Y., Andrews, R. M., … Thambisetty, M. (2018). Brain and blood metabolite signatures of pathology and progression in Alzheimer disease: A targeted metabolomics study. *PLoS Medicine*, *15*(1). https://doi.org/10.1371/journal.pmed.1002482

Verghese, P. B., Castellano, J. M., & Holtzman, D. M. (2011). Roles of Apolipoprotein E in Alzheimer's Disease and Other Neurological Disorders. *Lancet Neurology*, *10*. https://doi.org/10.1038/jid.2014.371

Vergoossen, L. W. M., Jansen, J. F. A., Backes, W. H., & Schram, M. T. (2020). Cardiometabolic determinants of early and advanced brain alterations: Insights from conventional and novel MRI techniques. *Neuroscience and Biobehavioral Reviews*, *115*(February), 308–320. https://doi.org/10.1016/j.neubiorev.2020.04.001

Vernooij, M. W., & Smits, M. (2012). Structural Neuroimaging in Aging and Alzheimer's Disease. *Neuroimaging Clinics of North America*, *22*(1), 33–55. https://doi.org/10.1016/j.nic.2011.11.007

Vilhjálmsson, B. J., Yang, J., Finucane, H. K., Gusev, A., Lindström, S., Ripke, S., … Price, A. L. (2015). Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *American Journal of Human Genetics*, *97*(4), 576–592. https://doi.org/10.1016/j.ajhg.2015.09.001

Viña, J., & Lloret, A. (2010). Why women have more Alzheimer's disease than men: Gender and mitochondrial toxicity of amyloid-ß peptide. *Journal of Alzheimer's Disease*, *20*(SUPPL.2), 527–533. https://doi.org/10.3233/JAD-2010-100501

Wand, H., Lambert, S. A., Tamburro, C., Iacocca, M. A., O'Sullivan, J. W., Sillari, C., … Wojcik, G. L. (2021). Improving reporting standards for polygenic scores in risk prediction studies. *Nature*, *591*(7849), 211–219. https://doi.org/10.1038/s41586-021-03243-6

Wang, R., Fratiglioni, L., Laukka, E. J., Lövdén, M., Kalpouzos, G., Keller, L., … Qiu, C. (2015). Effects of vascular risk factors and APOE ε4 on white matter integrity and cognitive decline. *Neurology*, *84*(11), 1128–1135. https://doi.org/10.1212/WNL.0000000000001379

Wardlaw, J. M., Valdés Hernández, M. C., & Muñoz-Maniega, S. (2015). What are white matter hyperintensities made of? Relevance to vascular cognitive

impairment. *Journal of the American Heart Association*, *4*(6), 001140. https://doi.org/10.1161/JAHA.114.001140

Wardlaw, Benvenisteb, & Nedergaard. (2020). Perivascular Spaces in the Brain: Anatomy, Physiology, and Contributions to Pathology of Brain Diseases. *Nature Reviews Neurology Review*.

Welsh, P., Welsh, C., Celis-Morales, C. A. ., Brown, R., Ferguson, L. ., Gray, S., … Sattar, N. (2020). Lipoprotein(a) and cardiovascular disease: prediction, attributable risk fraction and estimating benefits from novel interventions. *European Heart Journal*, *41*(Supplement_2), 1–28. https://doi.org/10.1093/ehjci/ehaa946.2833

Wheater, E., Shenkin, S. D., Muñoz Maniega, S., Valdés Hernández, M., Wardlaw, J. M., Deary, I. J., … Cox, S. R. (2021). Birth weight is associated with brain tissue volumes seven decades later but not with MRI markers of brain ageing. *NeuroImage: Clinical*, *31*(July). https://doi.org/10.1016/j.nicl.2021.102776

Wilkinson, T., Schnier, C., Bush, K., Rannikmäe, K., Henshall, D. E., Lerpiniere, C., … Sudlow, C. L. M. (2019). Identifying dementia outcomes in UK Biobank: a validation study of primary care, hospital admissions and mortality data. European Journal of Epidemiology, 34(6), 557–565. https://doi.org/10.1007/s10654-019-00499-1

Williams, S. M., & Haines, J. L. (2011). Correcting Away the Hidden Heritability. In Annals of Human Genetics (Vol. 75, Issue 3, pp. 348–350). https://doi.org/10.1111/j.1469-1809.2011.00640.x

Wilson, D. P., Jacobson, T. A., Jones, P. H., Koschinsky, M. L., McNeal, C. J., Nordestgaard, B. G., & Orringer, C. E. (2019). Use of Lipoprotein(a) in clinical practice: A biomarker whose time has come. A scientific statement from the National Lipid Association. *Journal of Clinical Lipidology*, *13*(3), 374–392. https://doi.org/10.1016/j.jacl.2019.04.010

Wilson, R. S., Leurgans, S. E., Boyle, P. A., & Bennett, D. A. (2011). Cognitive decline in prodromal alzheimer disease and mild cognitive impairment. *Archives of Neurology*, *68*(3), 351–356. https://doi.org/10.1001/archneurol.2011.31

Wojsiat, J., Laskowska-Kaszub, K., Mietelska-Porowska, A., & Wojda, U. (2017). Search for Alzheimer's disease biomarkers in blood cells: Hypotheses-driven approach. *Biomarkers in Medicine*, *11*(10), 917–931. https://doi.org/10.2217/bmm-2017-0041

Wu, A., Sharrett, A. R., Gottesman, R. F., Power, M. C., Mosley, T. H., Jack, C. R., … Coresh, J. (2019). Association of brain magnetic resonance imaging signs with cognitive outcomes in persons with nonimpaired cognition and mild Cognitive Impairment. *JAMA Network Open*, Vol. 2. https://doi.org/10.1001/jamanetworkopen.2019.3359

Xiao, E., Chen, Q., Goldman, A. L., Tan, H. Y., Healy, K., Zoltick, B., … Mattay, V. S. (2017). Late-Onset Alzheimer's Disease Polygenic Risk Profile Score Predicts Hippocampal Function. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *2*(8), 673–679. https://doi.org/10.1016/j.bpsc.2017.08.004

Yahagi, K., Kolodgie, F. D., Lutter, C., Mori, H., Romero, M. E., Finn, A. V., & Virmani, R. (2017). Pathology of human coronary and carotid artery atherosclerosis and vascular calcification in diabetes mellitus. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *37*(2), 191–204. https://doi.org/10.1161/ATVBAHA.116.306256

Yarchoan, M., Louneva, N., Xie, S. X., Swenson, F. J., Hu, W., Soares, H., … Arnold, S. E. (2013). Association of plasma C-reactive protein levels with the diagnosis of Alzheimer's disease. *Journal of the Neurological Sciences*, *333*(1–2), 9–12. https://doi.org/10.1016/j.jns.2013.05.028

Yusuf, Z. I., Dongarwar, D., Yusuf, R. A., Bell, M., Harris, T., & Salihu, H. M. (2019). Social Determinants of Overweight and Obesity Among Children in the United States. *International Journal of Maternal and Child Health and AIDS (IJMA)*, *9*(1), 22–33. https://doi.org/10.21106/ijma.337

Zachariah, J. P., Rong, J., Larson, M. G., Hamburg, N. M., Benjamin, E. J., Vasan, R. S., & Mitchell, G. F. (2018). Metabolic predictors of change in vascular function: Prospective associations from a community-based cohort.

Hypertension, *71*(2), 237–242. https://doi.org/10.1161/HYPERTENSIONAHA.117.10054

Zachariah, J. P., Rong, J., Larson, M. G., Hamburg, N. M., Benjamin, E. J., Vasan, R. S., & Mitchell, G. F. (2018). Metabolic predictors of change in vascular function: Prospective associations from a community-based cohort. *Hypertension*, *71*(2), 237–242. https://doi.org/10.1161/HYPERTENSIONAHA.117.10054

Zetterberg, H. (2019). Blood-based biomarkers for Alzheimer's disease—An update. *Journal of Neuroscience Methods*, *319*(October 2018), 2–6. https://doi.org/10.1016/j.jneumeth.2018.10.025

Zetterberg, H., & Blennow, K. (2021, December 1). Moving fluid biomarkers for Alzheimer's disease from research tools to routine clinical diagnostics. *Molecular Neurodegeneration*, Vol. 16. https://doi.org/10.1186/s13024-021-00430-x

Zhang, D., Tang, X., Shen, P., Si, Y., Liu, X., Xu, Z., … Gao, P. (2019). Multimorbidity of cardiometabolic diseases: Prevalence and risk for mortality from one million Chinese adults in a longitudinal cohort study. *BMJ Open*, *9*(3). https://doi.org/10.1136/bmjopen-2018-024476

Zhao, N., Liu, C. C., Qiao, W., & Bu, G. (2018). Apolipoprotein E, Receptors, and Modulation of Alzheimer's Disease. *Biological Psychiatry*, *83*(4), 347–357. https://doi.org/10.1016/j.biopsych.2017.03.003

Zhao, N., Ren, Y., Yamazaki, Y., Qiao, W., Li, F., Felton, L. M., … Bu, G. (2020). Alzheimer's Risk Factors Age, APOE Genotype, and Sex Drive Distinct Molecular Pathways. *Neuron*, 1–16. https://doi.org/10.1016/j.neuron.2020.02.034

Zheng, J., Haberland, V., Baird, D., Walker, V., Haycock, P. C., Hurle, M. R., … Gaunt, T. R. (2020). Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nature Genetics*, *52*(10), 1122–1131. https://doi.org/10.1038/s41588-020-0682-6

Zhong, X. M., Hou, L., Luo, X. N., Shi, H. S., Hu, G. Y., He, H. B., … Ning, Y. P. (2013). Alterations of CSF Cystatin C Levels and Their Correlations with CSF Aβ40 and Aβ42 Levels in Patients with Alzheimer's Disease, Dementia with Lewy Bodies and the Atrophic Form of General Paresis. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0055328