Sex Differences in White Matter Microstructure and Cognition in Healthy Aging

by

Lisa Ohlhauser
B.Sc. (Hons.), University of British Columbia, 2014
M.Sc., University of Victoria, 2018

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

University of Victoria

We acknowledge and respect the lək̓ʷəŋən peoples on whose traditional territory the university stands and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Supervisory Committee

Sex Differences in White Matter Microstructure and Cognition in Healthy Aging

by

Lisa Ohlhauser

B.Sc. (Hons.), University of British Columbia, 2014
M.Sc., University of Victoria, 2018

Supervisory Committee

Dr. Jodie Gawryluk, Department of Psychology
Supervisor

Dr. Stuart MacDonald, Department of Psychology
Departmental Member

Dr. Sandra Hundza, Department of Exercise Science, Physical & Health Education
Outside Member

Abstract

As the global population of older adults increases, it is crucial to study the healthy aging brain. Sex and gender are important determinants of health that may impact the aging trajectory, especially since women outlive men in most populations worldwide. To date, most studies examining sex and gender differences in aging have been cross-sectional and have focused on grey matter. In contrast, studies investigating age-related changes in white matter (WM) with longitudinal designs have been limited. The following three studies used available data from the healthy control group of the Parkinson Progression Markers Initiative (PPMI) to investigate biological sex differences in WM and cognition, as well as changes in these variables over time. In Study 1, sex differences in WM as measured by diffusion tensor imaging (DTI) were examined in a sample of 40 healthy older adults using tract-based spatial statistics. Results showed no significant differences in the most used DTI metrics of fractional anisotropy (FA) or mean diffusivity (MD) when controlling for total brain volume. In Study 2, longitudinal changes in WM were examined in the same healthy older adult sample over a one-year time frame. Stability in both FA and MD were found when examining changes in WM for males and females separately. Findings from Study 2 demonstrate that healthy older adult males and females showed similar stability in WM over a one-year period. In Study 3, longitudinal changes in cognition were examined in 193 healthy control participants from the PPMI. Linear mixed modeling was used to examine within-person change over time and between-person differences in cognition for up to six annual time points. Findings from Study 3 demonstrated that although males and females may have baseline differences in some cognitive domains, both sexes had similar stability in cognition over the five-year study period. Overall, these studies contribute to a growing body of literature examining sex similarities and differences in brain structure and function throughout the normal aging process. This research ultimately aids in the long-term goals of aging research to prevent disease and promote health and well-being into the latter decades of life.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables viii
List of Figures ix
Acknowledgments x
Dedication xi
Chapter 1: A Review of Sex and Gender Differences in the Aging Brain 1
  What is Healthy Aging? 2
  What are the Challenges in Sex and Gender Research? 5
  Why is Research in Sex and Gender in Healthy Aging Important? 9
  What are the Sex and Gender Differences in the Structure of the Healthy Adult Brain? 12
  What is Diffusion Tensor Imaging? 17
  What are the Sex and Gender Differences in White Matter Microstructure? 21
  What are the Sex and Gender Differences in Changes in White Matter Microstructure? 22
  What are the Sex and Gender Differences in Cognition? 23
  What are the Sex and Gender Differences in Cognitive Aging? 25
  Research Aims and Questions 29
    Study 1: Sex Differences in White Matter Microstructure in Healthy Older Adults 31
    Study 2: Longitudinal Changes in White Matter Microstructure in Healthy Older Adults 31
    Study 3: Sex Similarities and Differences in Cognition: A Longitudinal Study of Healthy Control Participants from the Parkinson's Progression Markers Initiative . 32
Chapter 2: Study 1 - Sex Differences in White Matter Microstructure in Healthy Older Adults 33
  Abstract 34
  Introduction 35
  Method 39
    Participants 39
    Diffusion-Weighted Images 40
    Procedure and Data Analysis 40
    Image Preprocessing 41
    Tract-based Spatial Statistics 41
    Voxelwise Statistical Comparisons 42
  Results 42
  Discussion 43
Chapter 3: Study 2 - Longitudinal Changes in White Matter Microstructure in Healthy Older Adults 47
  Abstract 48
  Introduction 50
  Methods 54
    Participants 54

    - Diffusion-Weighted Images 55
    - Procedure and Data Analysis 55
    - Image Preprocessing 55
    - Tract-based Spatial Statistics 56
    - Voxelwise Statistical Comparisons 56
  - Results 58
  - Discussion 64
- Chapter 4: Study 3 - Sex Similarities and Differences in Cognition: A Longitudinal Study of Healthy Control Participants from the Parkinson Progression Markers Initiative 68
  - Abstract 69
  - Introduction 71
  - Methods 75
    - Participants 75
    - Measures 76
    - Study Design 80
    - Model Specification and Equations 81
  - Procedure and Data Analysis 88
  - Results 88
    - Results for Question 1: What are the cognitive changes that occur among healthy adults over a five-year time frame (i.e., level-1, within-person change)? 89
    - Results for Questions 2: Are there sex differences in cognition (i.e., level-2, differences in intercepts)? 89
    - Results for Question 3: Are there sex differences in rates of change in cognition over five years (i.e., level-2, between-group differences in change, differences in slopes)? 90
  - Discussion 100
    - Finding 1: Stability in neuropsychological scores over time 100
    - Finding 2: Sex difference in neuropsychological scores at baseline 103
    - Finding 3: No sex difference in neuropsychological scores over time 105
    - Limitations 106
    - Strengths/Conclusion 107
    - Future directions 108
- Chapter 5: Conclusions and Future Directions 109
  - Dissertation Summary and Implication of Findings 109
  - Strengths 113
  - Limitations 116
  - Future Directions 120
  - Conclusion 123
- References 125
- Appendix A. Initial Baseline Inclusion/Exclusion Criteria for Healthy Control Participants from the Parkinson Progression Marker's Initiative 153
- Appendix B. Study 3 Syntax for Model Specification in R using lme4 154
- Appendix C. Study 3 Description of Final Model Fit and Specification for Each Neuropsychological Outcome Measure 157
  - Montreal Cognitive Assessment 157
  - Semantic Fluency Test 161

  - Symbol Digit Modalities Test 163
  - Benton Judgment of Line Orientation 165
  - Hopkins Verbal Learning Test – Revised (Immediate Recall) 167
  - Hopkins Verbal Learning Test – Revised (Delayed Recall) 170

Appendix D. Data Availability Statement and Funding Details 172

List of Tables

Table 1.1 *Definitions of Sex, Gender, and Sex/Gender* 7

Table 1.2 *Diffusion Tensor Imaging Metrics* 20

Table 1.3 *Longitudinal Studies of Cognitive Aging in Healthy Older Adults* 27

Table 2.1 *Participant Demographics* 43

Table 3.1 *Descriptive Statistics of Participants Selected from the PPMI Healthy Control Cohort* 59

Table 4.1 *Description of Neuropsychological Measures and Cognitive Domains Assessed* 79

Table 4.2 *Multi-level Model Specification Equations to Predict Neuropsychological Outcome Measures* 87

Table 4.3 *Participant Demographics and Descriptive Statistics* 91

Table 4.4 *Summary of the Multi-level Models to Predict Cognition on Seven Difference Neuropsychological Outcome Measures Over Five Years* 92

Table 4.5 *Results of Multilevel Models to Predict Montreal Cognitive Assessment Scores*.....*158*

Table 4.6 *Results of Multilevel Models to Predict Letter Number Sequencing Scores...160*

Table 4.7 *Results of Multilevel Models to Predict Semantic Fluency Scores*.....*162*

Table 4.8 *Results of Multilevel Models to Predict Symbol Digit Modalities Scores*.....*164*

Table 4.9 *Results of Multilevel Models to Predict Benton Judgment of Line Orientation Scores*.....*166*

Table 4.10 *Results of Multilevel Models to Predict Hopkins Verbal Learning Test - Revised Immediate Recall Scores*.....*169*

Table 4.11 *Results of Multilevel Models to Predict Hopkins Verbal Learning Test - Revised Delayed Recall Scores*.....*171*

List of Figures

Figure 1.1 The life course model of the Scaffolding Theory of Aging - Revised (STAC-r). 5

Figure 1.2 Diffusion tensor imaging model. 19

Figure 3.1 Decreases in MD (blue) from baseline to follow-up for all participants (p<.05). 60

Figure 3.2 Changes in whole brain FA values for all participants from baseline to follow-up separated by sex. 61

Figure 3.3 Changes in whole brain MD values for all participants from baseline to follow-up separated by sex. 62

Figure 3.4 Changes in Montreal Cognitive Assessment Scores for females and males from baseline to follow-up. 63

Figure 4.1 Schematic of the 2-level linear mixed model. 80

Figure 4.2 Montreal Cognitive Assessment Scores over time. 93

Figure 4.3 Letter Number Sequencing Scores over time. 94

Figure 4.4 Semantic Fluency Scores over time. 95

Figure 4.5 Symbol Digit Modalities Scores over time. 96

Figure 4.6 Benton Judgment of Line Orientation Scores over time. 97

Figure 4.7 Hopkins Verbal Learning Test (Revised) Immediate Recall Scores over time. 98

Figure 4.8. Hopkins Verbal Learning Test (Revised) Delayed Recall Scores over time. 99

Acknowledgments

I express my sincerest gratitude to all those who made this dissertation possible. Foremost, it is an honor thank my supervisor, Dr. Jodie Gawryluk, for her continuous support and mentorship over the past seven years. This work would not have been possible without her encouragement and guidance. I am also deeply grateful to my supervisorial committee members, Drs. Stuart MacDonald and Sandra Hundza for their helpful suggestions and guidance for the study design and statistical analysis.

I thank my lab members for their continued support and friendship. I am also grateful to all the professors, students, and staff in the Clinical Psychology program at the University of Victoria. I have been fortunate to be part of this academic community.

I am sincerely grateful to my partner and my family for their endless love and support during throughout my doctoral training.

I express my gratitude to the Canadian Institute of Health Research and the Women, Sex, Gender and Dementia & Canadian Consortium on Neurodegeneration and Aging (CCNA) Team 9 for funding this research. Finally, thank you to the Parkinson's Progression Markers Initiative for sharing their data with myself and other researchers around the world.

Dedication

For DW.

Chapter 1: A Review of Sex and Gender Differences in the Aging Brain

It is vital to understand changes in the aging brain to promote health and longevity into the later decades of life as the global population of older adults is increasing. Worldwide, the population of adults over the age of 65 is expected to double from 703 million to 1.5 billion by 2050 and the population of adults over the age of 80 is expected to triple from 143 million to 426 million by 2050 (United Nations et al., 2020). In recognition of the fact that populations around the globe are aging at a faster pace than ever before, the United Nations has declared 2021-2030 the Decade of Healthy Aging, which outlines a 10-year collaborative plan to improve the lives of older adults, their families, and communities around the world. Notably, age is a primary risk factor for most types of dementia (Hou et al., 2019), including Alzheimer's disease (Reitz & Mayeux, 2014), Lewy Body Dementia (Jellinger, 2003), and Parkinson's disease (Reeve et al., 2014). Unfortunately, the number of individuals with dementia is also expected to rise significantly with the growing aging population.

A great deal of research has focused on disease specific changes in the brain. However, the expected structural and functional changes that occur as part of normal healthy aging are not yet entirely understood. Studying the expected changes in healthy aging could provide important information about the etiology and prevention of age-related diseases. More specifically, modelling healthy aging could be used to differentiate the early subtle changes in brain structure and function that may signal the beginning of a disease process and a point in time where prevention of decline and promotion of successful aging would ideally begin. A review of the literature is discussed next.

What is Healthy Aging?

Many terms have been used to describe the aging process, including “normal aging”, “typical aging”, “successful aging”, and “optimal aging”. One of the most used terms is “healthy aging”, which is defined by the World Health Organization as, “developing and maintaining the functional ability that enables well-being in older age.” Functional ability encompasses the intrinsic capability of an individual, the environment in which they live, and the interactions among them.

Another important component of aging is “cognitive aging”, which describes the process of gradual, ongoing, and highly variable changes in cognitive functions that occur as people get older (Institute of Medicine (U.S.) et al., 2015). Cognitive aging occurs in all individuals as they age and regardless of initial cognitive ability (Park & Reuter-Lorenz, 2009). Cognition is multidimensional and captures many specific mental abilities that are fundamental to engagement in everyday life. Some of the most conceptualized domains are attention, memory, executive function, visuospatial abilities, and processing speed.

Importantly, specific cognitive abilities may not necessarily change uniformly with increasing age. Some cognitive abilities, such as vocabulary, are resilient to brain aging and may even improve with age. Other cognitive abilities, such as conceptual reasoning, memory, and processing speed, decline gradually over time (Salthouse, 2010). There is also interindividual and intraindividual variability in cognitive aging. Some individuals may experience relative stability in some domains, while others may decline. In some cases, there is potential for improvements in cognitive abilities over time (Harada et al., 2013).

Biological and environmental factors that influence the structure and function of the brain are thought to contribute to the cognitive aging process. Several models, such as the Scaffolding Theory of Aging and Cognition (STAC) (Park & Reuter-Lorenz, 2009) and its revised version (STAC-r) (Reuter-Lorenz & Park, 2014) have been developed to better understand the factors that contribute to changes in brain structure, function, and cognition that occur during healthy aging. A primary component of the original STAC model is the concept of compensatory scaffolding. This concept describes how engagement of supplementary neural circuitry can provide additional computational support to preserve cognitive function in the presence of adverse changes in brain structure (Reuter-Lorenz & Park, 2014). Thus, both positive and negative neural processes contribute to overall cognitive function and changes in cognition that occur over time.

The concept of compensatory scaffolding is retained in the STAC-r model and two new constructs which are influenced by life course experience were added: neural resource enrichment and neural resources depletion. Neural resource enrichment includes ongoing factors that contribute to enhancements in brain structure and function. Examples of potential neural resource enrichment factors include intellectual engagement, education, physical and cardiovascular fitness, multilingualism, and increased engagement in leisure activities, which have all been associated with beneficial cognitive outcomes (Reuter-Lorenz & Park, 2014). It is proposed that neural resource enrichment factors can either directly enhance or preserve brain function through structural and functional changes in the brain (e.g., promoting efficient connectivity, increasing cortical thickness or synaptic density) or indirectly enhance or preserve brain function through increased compensatory scaffolding. Neural resource depletion includes factors that negatively contribute to brain structure and function. Examples of potential neural depletion factors include the presence of APOE e4 alleles, stress, vascular disease, low socioeconomic status, depression, neuroticism, and acquired brain injury (e.g., head trauma, toxin exposure), which have all been associated with poorer cognitive outcomes (Reuter-Lorenz & Park, 2014). The overall STAC-r model posits that these neural resource enrichment and depletion factors impact the structure and function of the brain to influence cognition through compensatory scaffolding (see *Figure 1.1*). Although many factors have been proposed to influence cognitive aging, sex and gender have not yet been well examined or integrated into aging models such as STAC-r. It is important to recognize that there is a growing consensus that the terms sex and gender are not equivalent, despite many researchers using these terms interchangeably. The challenges of sex and gender research will be discussed next.

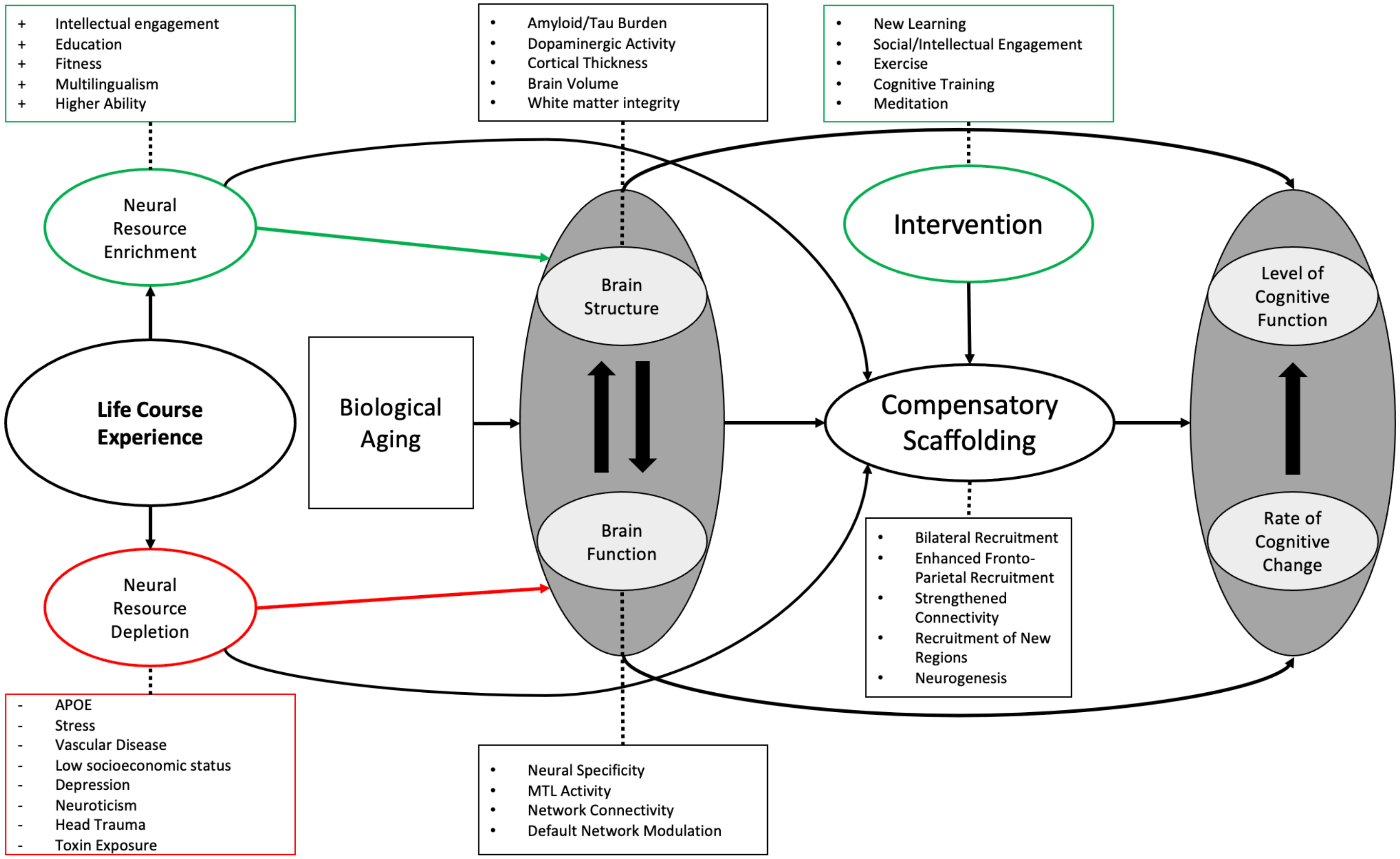

Figure 1.1 The life course model of the Scaffolding Theory of Aging - Revised (STAC-r).
The STAC-r was informed by findings from neuroimaging research and posits that neural resource enhancement and depletion factors throughout the life course can impact the structure and function of the brain and overall cognitive function through compensatory scaffolding. Figure adapted from Reuter-Lorenz & Park, 2014.

What are the Challenges in Sex and Gender Research?

Sex and gender research intersects many disciplines, including psychology (Hyde et al., 2019), neuroscience (Jordan-Young & Rumiati, 2012), and neuroethics (Roy, 2012). Inconsistent terminology and difficulties applying these concepts in research still exist (Day et al., 2016). Several national and international organizations have reached a general consensus on the definition of sex and gender, which are provided in Table 1. These definitions generally state that *sex* refers to biological and physical characteristics of humans whereas *gender* refers to socially constructed characteristics of personal identity. Notably, separating the variables of sex and gender in research is difficult. There is a bidirectional entanglement of these factors whereby it is extremely challenging to differentiate inborn sex factors from social and environmental gender learning (Eliot et al., 2021). Several researchers have suggested that the composite term “sex/gender” be used (Kaiser, 2012; Rippon et al., 2014). Given the inconsistency in terminology throughout the literature, it is agreed that the term “sex/gender” will be used, unless the term sex or gender is clearly identified as the variable of interest with a stated meaning that agrees with the consensus definitions outlined in Table 1.1.

Table 1.1 *Definitions of Sex, Gender, and Sex/Gender*

Organization	Sex	Gender	Sex/Gender
Canadian Institute of Health Research[1]	"Refers to a set of **biological** attributes in humans and animals. It is primarily associated with physical and physiological features including chromosomes, gene expression, hormone levels, and reproductive/sexual anatomy. Sex is usually categorized as female or male but there is variation in the biological attributes that comprise sex and how those attributes are expressed."	"Refers to the **socially constructed** roles, behaviours, expressions and identities of girls, women, boys, men, and gender diverse people. It influences how people perceive themselves and each other, how they act and interact, and the distribution of power and resources in society. Gender identity is not confined to a binary (girl/woman, boy/man) nor is it static; it exists along a continuum and can change over time. There is considerable diversity in how individuals and groups understand, experience and express gender through the roles they take on, the expectations placed on them, relations with others and the complex ways that gender is institutionalized in society."	
World Health Organization[2]	"Refers to the different **biological** and physiological characteristics of females, males and intersex persons, such as chromosomes, hormones and reproductive organs."	"Refers to the characteristics of women, men, girls and boys that are **socially constructed.** This includes norms, behaviours and roles associated with being a woman, man, girl or boy, as well as relationships with each other. As a social construct, gender varies from society to society and can change over time."	
National Institute of Health[3]	"Refers to **biological** differences between females and males, including chromosomes, sex organs, and endogenous hormonal profiles."	"Refers to **socially constructed** and enacted roles and behaviors which occur in a historical and cultural context and vary across societies and over time."	
Kaiser, 2012			"It is impossible for the variables of ''sex'' and ''gender'' to be categorized as solely biological or

	solely social components or to be measured or recorded as such, and for this reason, they should in fact be regarded as a unity and designated as sex/gender."
Rippon et al., 2014	"Neural phenotypes represent the complex entanglement of biological and environmental factors, such that it is generally not possible to entirely isolate the two. Thus, we use the composite term “sex/gender” as a way to refer to this irreducible complexity"

---

[1] https://cihr-irsc.gc.ca/e/48642.html
[2] https://www.who.int/health-topics/gender
[3] https://orwh.od.nih.gov/sex-gender

Systematic barriers in sex and gender research have been recognized in biomedical research (Mazure & Jones, 2015). To encourage the inclusion of these variables in research in Canada, the Canadian Institutes of Health Research (CIHR) introduced a new policy in 2010. All applicants for CIHR funding opportunities were to indicate whether their research designs accounted for sex or gender. Although there was an overall increase in the percentage of researchers who reported that they would incorporate sex or gender into their research, biomedical researchers were least likely to do so compared with fellow researchers in clinical, health systems, or population health (Johnson et al., 2014). Further, these data are based on self-reported intent to investigate sex and gender as stated on funding applications, which does not necessarily reflect whether these investigations were conducted or published. There are also disparities in CIHR research funding. Grant applications in sex and gender-based research in neuroscience and cellular processes were less likely to be awarded than sex and gender-based research in other disciplines (Johnson et al., 2014). Despite these challenges, there are significant efforts by CIHR to promote and support sex and gender-based research, including several online training modules from the Institute of Gender and Health found here: https://cihr-irsc.gc.ca/e/49347.html.

There are several well-known differences among males and females that highlight the need for sex and gender-based health research, especially with respect to sex and gender differences in aging, which are discussed next.

Why is Research in Sex and Gender in Healthy Aging Important?

Sex and gender are important determinants of health that may significantly impact the aging trajectory (Banister et al., 2012). There are notable sex differences in life expectancy and the prevalence of specific diseases throughout the lifespan. Women outlive men in virtually all populations globally (Barford et al., 2006). According to the World Health Organization (2020), the average life expectancy is 74.2 years for women and 69.8 years for men. Canadian life expectancies are higher than the global average, but still maintain significant sex/gender differences with life expectancy at 84.7 years for women and 80.9 years for men (World Health Organization, 2020). There are also sex and gender differences in the prevalence of many disorders of the central nervous system across the lifespan. Neurodevelopmental disorders, including Autism Spectrum Disorder (Loomes et al., 2017), Attention-Deficit/Hyperactivity Disorder (Polanczyk et al., 2007), and dyslexia (Rutter et al., 2004) are more prevalent in males, while many psychiatric disorders including anxiety disorders (McLean et al., 2011), depression (Salk et al., 2017), and eating disorders (Hudson et al., 2007), are more prevalent in females. Sex and gender differences persist in the latter decades of life and there are significant sex and gender differences in many diseases associated with aging. Alzheimer's disease and Multiple Sclerosis are more prevalent in women, while Parkinson's disease and Amyotrophic Lateral Sclerosis are more prevalent in men (Zagni et al., 2016).

There are several factors that may account for sex and gender differences in aging. From a sociocultural perspective, males and females experience different societal expectations that influence life experiences and behaviour that shape the neural resource enhancement and enrichment factors described in the STAC-r model (see *Figure 1.1*). Gender identity is inherently linked with factors that determine cognitive reserve, such as educational attainment, occupation, diet, exercise, substance use, and stress. For example, women historically had fewer opportunities for education and enriching employment compared to men (Weber et al., 2014). This reduced the opportunity for many women in older cohorts to build cognitive reserve. Lifestyle factors, including diet and exercise, can also be influenced by gender. For example, men may have more opportunities to participate in sports or physically demanding jobs, which could significantly influence overall cardiovascular health (Grundy & Sloggett, 2003). There are also gender differences in substance use, including alcohol and cigarette use, which have historically been more common in men, which increase the risk of health morbidities and premature mortality (McHugh et al., 2018). With respect to Alzheimer's disease and other dementias, some have suggested that there may be a survivorship bias for women, such that women live longer than men, increasing the opportunity for Alzheimer's disease and other age-related diseases to develop in women (Andrew & Tierney, 2018).

From a biological perspective, men and women have different sex chromosomes and hormones that can influence overall health and aging trajectories. Estrogen is a primary sex hormone in women and is thought to be neuroprotective in premenopausal women. Estrogen has been found to promote the growth and survival of cholinergic neurons, increases cholinergic activity, and promote the nonamyloidogenic metabolism of the amyloid precursor protein (Janicki & Schupf, 2010). Some research also suggests that some forms of estrogen may have positive effects on cognition and neuroplasticity. For example, 17β–estradiol may facilitate cell signalling in certain pathways by increasing dendritic spine density, which may lead to increase synaptic plasticity and enhanced memory (Galea et al., 2017). Sex hormones change significantly throughout the lifespan and there is a significant drop in estrogen in women during menopause. As older men continue to produce some levels of testosterone that is converted to estradiol in the brain, serum levels of estrogen may actually be higher in older men than postmenopausal women (Janicki & Schupf, 2010).

Some cardiovascular risk factors are specifically associated with the female reproductive system. Reproductive endocrine disorders, such as polycystic ovary syndrome and primary ovarian insufficiency, as well as complications during pregnancy, such as pregnancy-induced hypertension, pre-eclampsia, gestational diabetes, and preterm birth, are associated with increased risk for cardiovascular disease, coronary heart disease, hypertension, stroke, and Type II diabetes (Gao et al., 2019). Genetic and epigenetic factors that can also affect men and women differently (Ngun et al., 2011) and research investigating sex-specific genes influencing longevity (Zeng et al., 2018) and heritability of specific disease biomarkers, such APOE in Alzheimer's disease, are underway (H. Wang et al., 2021). Finally, innate and adaptive immune responses differ between males and females at younger and older ages, which can also influence the propensity for inflammation and disease (Gubbels Bupp et al., 2018).

In summary, biological sex differences and gendered experiences are clearly important factors that influence overall health throughout the lifespan. It is especially crucial to investigate brain structure and function in older healthy adults so that sex and gender-specific neural enhancement and depletion factors may be elucidated. The known sex and gender differences in cognition are discussed next.

What are the Sex and Gender Differences in the Structure of the Healthy Adult Brain?

As described in the STAC-r model, neural resource enrichment and depletion factors are thought to impact the structure and function of the brain. Through neuroimaging methods, the structure and function of the brain can be assessed non-invasively and used as a biomarker for cognitive aging. Magnetic resonance imaging (MRI) is based on the principle of nuclear magnetic resonance, which capitalizes on the abundance of hydrogen atoms (i.e., single protons) largely present in human tissue in water molecules. This method has been used frequently to study the brain. A comprehensive synthesis of the literature conducted by Eliot and colleagues (2021) concluded that the overall sex/gender differences in brain structure and function are very small after accounting for the significant differences in overall brain volume (Eliot et al., 2021).

On average males are taller, weigh more, and have larger heads than females. This trend continues neuroanatomically, where males have larger brain sizes on average than females. This is most often measured as total intracranial volume (ICV), which includes all of the anatomy within the skull, and total brain volume (TBV), which excludes the meninges and ventricles. One of the largest meta-analyses to date found that on average, males had 11% larger ICVs and 12% larger TBVs than females (Ruigrok et al., 2014). Males were also found to have significantly larger volumes in many respective parts of the brain than females, including 10% larger cerebrums, 9% more gray matter, 13% more white matter, 11.5% more cerebrospinal fluid, and 9% larger cerebellums (Ruigrok et al., 2014). It is important that the vast majority of studies included in the analysis were representative of individuals in the 18-59 age range, although age categories from across the lifespan were also included.

One of the largest single studies of sex/gender difference in brain structure completed to date included 5216 individuals from the ages of 44-77 from the UK Biobank (Ritchie et al., 2018). Similar results were found, where males have larger overall brain volumes (i.e., TBV) and larger volumes of both grey and white matter. Males also had larger volumes in all subcortical structures measured, including the hippocampus, nucleus accumbens, amygdala, caudate nucleus, pallidum, putamen, and thalamus. However, specific structures in the brain do not necessarily scale linearly with body height or the overall size of the brain. As such, it is unclear whether sex/gender differences are due to overall size or more specific biological processes underlying the structure and function of the brain (Jahanshad & Thompson, 2017). The methodology also varies across studies and there is inconsistency as to whether or not TBV or ICV is controlled.

Although males have larger total brain volumes than females, many specific regional asymmetries for males and females related to both volume and tissue density have been found. A meta-analysis of sex/gender comparisons of 16 studies involving voxel-based grey matter volume and nine studies involving voxel-based grey matter density, found that on average males have larger grey matter volumes and higher grey matter densities in the left amygdala, hippocampus, insular cortex, putamen and regions of the cerebellum (Ruigrok et al., 2014). Males were also found to have higher grey matter densities in the left claustrum, as well as larger grey matter volumes in the bilateral anterior parahippocampal gyri, posterior cingulate gyri, precuneus, temporal poles, areas in the left cingulate, and in the right amygdala. Conversely, females have been found to have higher grey matter tissue density in the left frontal pole, and larger grey matter volumes in the right frontal pole, inferior and middle frontal gyri, pars triangularis, planum temporale/parietal operculum, anterior cingulate gyrus, insular cortex, Heschl's gyrus, bilateral thalami and precuneus, left parahippocampal gyrus and lateral occipital cortex. However, these results come from just one meta-analysis and there was significant variability in sample size, age range, image acquisition parameters, and statistical models and thresholds to control for various factors such as differences in TBV (Ruigrok et al., 2014). A few meta-analyses and several other studies investigating sex/gender differences in subcortical volumes in various brain structures have been conducted. Importantly, there is no single subcortical structure that has reliably shown sex/gender differences across all studies. Although many studies have found males to have larger putamen and amygdala than females, these sex/gender differences are exceedingly small at 1-3% (Eliot et al., 2021). Overall, the evidence suggests that there are negligible sex/gender differences in subcortical structural volumes after controlling for overall brain size.

Similarly, results from neuroimaging studies that matched male and female participants by TBV or ICV have found that the majority of the cortex is the same regardless of sex/gender. However, results have been mixed and large studies in healthy older adults are limited. No sex/gender differences were found in a study examining cortical gray matter volumes in a sample of 42 males and females matched for ICV with a mean age of 75 years (Voevodskaya et al., 2014). Females were found to have larger cortical volumes than males in a study including 48 males and females with a mean age of 43.4 years matched by TBV (Luders et al., 2009). Females had larger volumes in the left orbitofrontal gyrus, superior temporal gyrus, and superior frontal gyrus. Both females and males had regions larger than the other sex/gender in two other studies, one of which included 70 males and 70 females with a mean age of 21.2 years matched for TBV (L. Wang et al., 2012) and another of which included 148 males and females with a mean age of 21.9 years matched for ICV (Sanchis-Segura et al., 2019). When accounting for TBV and ICV, sex/gender differences in cortical volumes are quite similar, but evidence suggests that small variations may exist depending on sample characteristics and study methodology.

Another brain region of interest has been cortical thickness. Eliot et al (2021) reviewed 25 highly cited studies with large samples sizes that examined sex/gender differences in cortical thickness. They concluded that there are no reliable sex/gender differences. Although many studies found larger cortical thicknesses in females than males, Eliot and colleagues argue that this may be due to methodological issues with a particular software program (i.e., FreeSurfer) as post-mortem histological studies have found slightly greater cortical thicknesses in males.

There have also been many studies investigating the inter-hemispheric connections in the brain. The corpus callosum is the largest inter-hemispheric structure and has garnered the most investigations in sex/gender differences in size. Several meta-analyses have been conducted and have found mixed results. Males were found to have larger corpus callosums in studies that did not control for overall brain size (Bishop & Wahlsten, 1997; Driesen & Raz, 1995), while females were found to have larger corpus callosums in studies that implemented methods to control for overall brain size (Driesen & Raz, 1995; Smith, 2005). Results from studies that matched male and female participants for ICV have also been mixed, with some finding larger corpus callosums in males (Sullivan et al., 2001), some finding larger corpus callosums in females (Ardekani et al., 2013; Shiino et al., 2017), and some finding no significant sex/gender differences (Luders et al., 2014). However, many of these studies had small samples sizes. The largest study of corpus callosum size to date comes from normative data of 2790 individuals from age 18-94 from 21 independent research groups which found significantly larger corpus callosum in females (Potvin et al., 2016, 2018). However, the effect size was very small, indicating that sex accounts for less than 1% of the variance in corpus callosum volume.

Overall, there are minimal sex/gender differences in the brain apart from larger brain volumes in males (Eliot et al., 2021). Most of the studies discussed thus far have been focused on grey matter, which describes the cell bodies within the brain. However, despite representing approximately 50% of the brain, studies of sex/gender differences in white matter microstructure have been limited. White matter comprises the neuronal axons that connect structures throughout the brain. Diffusion tensor imaging (DTI) is a magnetic resonance imaging (MRI) technique that uses the principals of water diffusion to infer the integrity of white matter microstructure throughout the brain. Even fewer studies have examined sex/gender differences in white matter and in the context of aging. A brief overview of DTI is provided next before a more tailored discussion of sex/gender differences in white matter microstructure and longitudinal changes in white matter in healthy adults.

What is Diffusion Tensor Imaging?

Diffusion tensor imaging (DTI) is a method to quantify the molecular diffusion of water molecules measured by diffusion-weighted imaging, a variant of conventional MRI. DTI relies on the application of magnetic field gradients in different directions to estimate the strength and direction of the diffusion of water within each voxel. This information can be used to infer in vivo information about specific properties of various structures and tissue within the brain and is particularly useful for estimating the microstructural integrity of white matter (i.e., axons) within the brain. Diffusion describes the random motion of molecules due to kinetic energy and can be described as “isotropic” or “anisotropic”.

Isotropic diffusion refers to unrestricted diffusion, which occurs when there are few structural barriers, such as in cell bodies (i.e., gray matter) and cerebrospinal fluid. Anisotropic diffusion refers to restricted diffusion, which occurs when physical barriers are present, such as along tubular structures such as axons (i.e., white matter).

The signal measured in a given voxel depends on the strength and duration (b-value) as well as the direction of gradients, and the local tissue microstructure (e.g., presence of barriers to diffusion, such as bundles of axons). If diffusion of protons has occurred in a certain direction, the signal will be attenuated, with more attenuation indicating more diffusion in the direction of the applied gradient. A minimum of six (and more typically 30) gradient directions are used to calculate the overall diffusion of water molecules within a voxel in an ellipsoid shape using a 3 x 3 symmetric matrix (i.e., a tensor model) (Basser et al., 1994). Each voxel contains three eigenvalues (λ_1, λ_2, λ_3) and their respective eigenvectors (V_1, V_2, V_3) which measure the magnitude of and direction of diffusion (see *Figure 1.2*).

To date, several DTI metrics using eigenvalues have been developed that provide information about the overall diffusion properties within each voxel. The most commonly reported metrics are fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD). Please see Table 1.2 for a brief description and common interpretation of each DTI metric. Fractional anisotropy provides an estimate of the overall direction of diffusion. In isotropic diffusion, all eigenvalues are equal in magnitude (i.e., $\lambda_1 = \lambda_2 = \lambda_3$) which results in an FA value of 0. In anisotropic diffusion, the eigenvalues are unequal which results in FA values closer to 1 (see *Figure 1.2*). Mean diffusivity is calculated by taking the mean of the three eigenvalues and describes the overall diffusion within a voxel. In general, decreased FA and increased MD suggest less restrictive diffusion and these patterns are often interpreted as reflecting the loss of physical barriers of diffusion and decreased microstructural integrity within white matter. However, the biophysical interpretations of DTI metrics are often complicated due to multiple crossing fibres, which can lead to Type I (i.e., false positive) or Type II (i.e., false negative) errors (Figley et al., 2022).

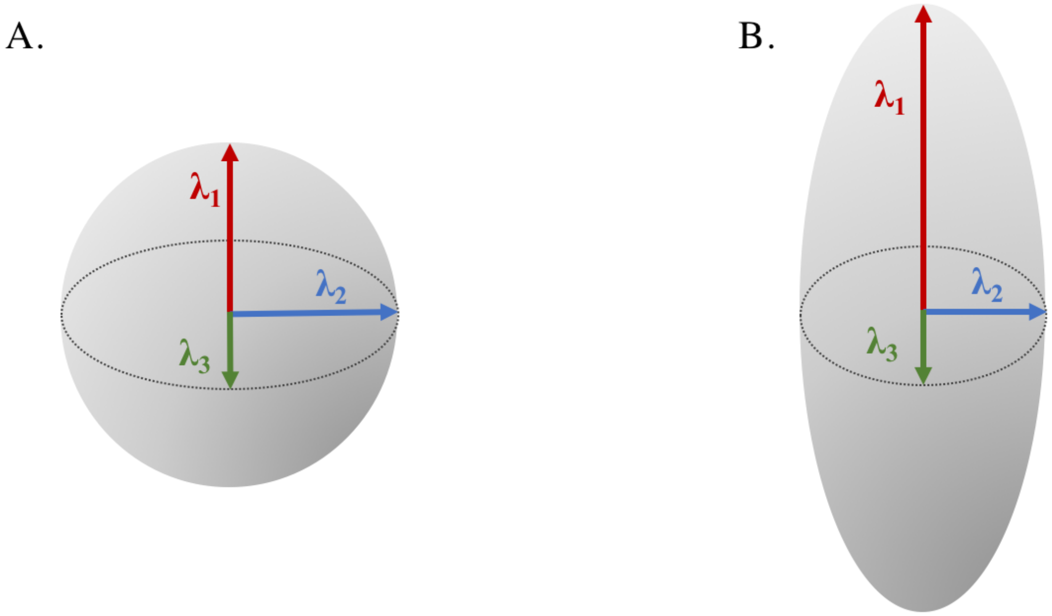

Figure 1.2 Diffusion tensor imaging model.

A. Isotropic diffusion, where eigenvalues are equal ($\lambda_1 = \lambda_2 = \lambda_3$), and B. Anisotropic diffusion, where eigenvalues are unequal.

Table 1.2 *Diffusion Tensor Imaging Metrics*

	Metric			
	Fractional Anisotropy	Mean Diffusivity	Radial Diffusivity	Axial Diffusivity
Abbreviation	FA	MD	RD	AD
Mathematic equation	$\sqrt{\frac{1}{2}\frac{\sqrt{(\lambda_1-\lambda_2)^2+(\lambda_2-\lambda_3)^2+(\lambda_3-\lambda_1)^2}}{\sqrt{\lambda_1{}^2+\lambda_2{}^2+\lambda_3{}^2}}}$	$\frac{\lambda_1+\lambda_2+\lambda_3}{3}$	$\frac{\lambda_2+\lambda_3}{2}$	λ_1
Description	Degree of elongation of the diffusion tensor ellipsoid	Average of the diffusivity values of the three axes of the diffusion ellipsoid	Average of the diffusivities in the axes perpendicular to the major axes of diffusion	Diffusivity in the direction of maximum diffusion in the voxel
Common Interpretation	Summary of microstructural integrity	Inverse measure of membrane density	Increased in white matter with demyelination	Variable with white matter changes
Direction typically associated with neuro-degeneration[1]	Decreases	Increases	Variable	Variable
Sensitive to:	Wide range of pathologies	Cellularity, edema, and necrosis	Myelin loss	Axonal injury
Limitations	Crossing fibres with high integrity can have low FA	Large variability in measurement; crossing fibers	Voxels with crossing fibres can increase RD	Voxels with crossing fibers can increase AD

Note. λ_1, λ_2, and λ_3 reflect the longest, middle, and shortest eigenvalues in a diffusion tensor imaging model (see *Figure 1.2*). [1]Decreased FA and increased MD associated with decreases in physical barriers to diffusion. Information adapted from Alexander, Lee, Lazar, & Field, 2007; Figley et al., 2022; and Van Hecke, Emsell, & Sunaert, 2016.

What are the Sex and Gender Differences in White Matter Microstructure?

Despite the brain being comprised of approximately 50% white matter, few studies thus far have examined sex differences using DTI metrics. DTI studies investigating sex/gender differences in white matter microstructure thus far have largely been cross-sectional and conducted on groups in early to mid-adulthood (i.e., 20-40 years). Several studies with participants in young to middle-aged adult groups have found mixed regional sex differences in DTI metrics. For example, Chou et al. (2011) found increased FA in the fronto-occipital fasciculus, body of the corpus callosum, and parahippocampal gyrus in females and increased FA in the bilateral internal capsule, medial frontal gyrus, fusiform gyrus, hippocampus, insula, postcentral gyrus, frontal and temporal lobe in males. Kanaan et al. (2014) corroborated the finding of increased FA in the corpus callosum in females and also found increased FA in cerebellar areas in males. In contrast, Menzler et al. (2011) found increased FA and decreased RD in the thalamus, cingulum, and corpus callosum in males compared to females.

One of the largest samples to date comes from a UK Biobank Study (Cox et al., 2016) that included 3513 healthy older adults from age 44.64 to 77.12 years. With respect to white matter microstructure, males were found to have overall higher FA than females, while females were found to have overall higher neurite orientation dispersion (a measure of spatial configuration of neurite structures), than males.

Taken together, no consistent sex/gender differences in white matter microstructure have been observed. Factors such as sample size, participant age, scanner type, analysis software and methods may contribute to the mixed findings. As discussed previously, it is important to control for overall brain size when conducting sex/gender-based analyses in the structure of the brain. Some have found that most, but not all, of the significant sex differences diminished after adjusting for total intracranial volume (Takao et al., 2014). However, others have found unequivocal sex/gender differences. For example, males have been found to have increased FA in the anterior genu region of the corpus callosum compared to females (Westerhausen et al., 2011).

What are the Sex and Gender Differences in Changes in White Matter Microstructure?

Changes in white matter microstructure are expected throughout the lifespan. A systematic review of 30 DTI studies has summarized the known changes in white matter in healthy individuals across the lifespan: white matter development shows a curvilinear pattern where there is a rapid increase in FA during the neonatal period, which continues into childhood and adolescence, before slowing and plateauing in mid-adulthood, then declining with old age (Yap et al., 2013). This has been corroborated by others who found that overall FA peaks in most tracts between the ages of 23 to 39 (Kochunov et al., 2012). In general, MD has shown the opposite trend (Yap et al., 2013). Regionally, it appears that peak FA is reached in associative white matter tracts in later life and shows more age-related decline than areas that reach peak FA at a younger age, such as the motor and sensory tracts (Kochunov et al., 2012; Yap et al., 2013). A “last-in-first-out” hypothesis of brain aging has emerged, which posits that the brain areas to develop last are the most vulnerable to degeneration (Raz et al., 2010; Raz & Rodrigue, 2006).

Several cross-sectional studies have also examined the effects of age on white matter microstructure in addition to sex. A particularly large study with 857 healthy subjects ranging in age from 24.9-84.8 years found that global FA negatively correlated with age, while AD and RD positively correlated with age (Inano et al., 2011). Regional sex differences in white matter microstructure were also found. Males had higher FA in the splenium of the corpus callosum, bilateral corona radiata, posterior limb of the internal capsule, cerebral peduncle, external capsule, bilateral cingulum, bilateral superior longitudinal fasciculus, and bilateral middle cerebellar peduncle. Females had higher FA in the column of the fornix. However, no specific regions had a significant interaction between age and sex (Inano et al., 2011). A similarly sized study of 831 participants aged 11-90 years modelled FA values throughout the lifespan and found significant age by sex interactions. Males had higher overall average FA, higher FA in the cingulum, corticospinal tract, and splenium of the corpus callosum, and showed a slightly higher rate of decline in FA than females during aging (Kochunov et al., 2012). Results from the UK Biobank Study found that older age was also related to poorer white matter microstructure as evidenced by lower FA and higher MD (Cox et al., 2016).

Although there are significant complexities when examining sex differences in brain anatomy (Jahanshad & Thompson, 2017), DTI studies have demonstrated that significant changes in white matter microstructure occur across the lifespan and these findings have contributed to theories about the structural brain changes that may occur in healthy aging.

What are the Sex and Gender Differences in Cognition?

There has been great interest and controversy in the study of sex and gender differences in cognition since the 1970s (Maccoby & Jacklin, 1974). Although significant sex/gender differences have been found in specific cognitive abilities, including visuospatial working memory (Voyer et al., 2017), mental rotation (Maeda & Yoon, 2013), episodic memory (Asperholm et al., 2019), and verbal abilities (Barel & Tzischinsky, 2018; Zhang et al., 2017), the effect sizes of these differences are usually quite small. The "gender similarities hypothesis" posits that males and females are largely similar on most, but not all variables (Hyde, 2005). In a seminal review (Hyde, 2014), a slight male advantage in spatial performance was acknowledged, but otherwise, males and females were argued to have highly similar cognitive abilities, even in mathematical and verbal abilities, which were previously thought to significantly differ (Maccoby & Jacklin, 1974). Although males and females are more similar cognitively than they are different, sex/gender differences in cognition remains an ongoing area of investigation and are relevant for understanding the cognitive changes that may occur in healthy aging.

There are small sex/gender differences in visuospatial working memory. A meta-analysis of 180 effects sizes from 98 samples found an overall small significant advantage in visuospatial working memory for males (Voyer et al., 2017). This pattern was consistent across all visuospatial working memory tasks, including Corsi blocks, n-back, memory for patterns and tokens. However, females were found to have a small but significant advantage over males on a memory for location task.

A meta-analysis found a medium to large effect size for sex differences in mental rotation as assessed by the Purdue Spatial Visualization Test – Visualizations by Rotation (PSVT-R) (Maeda & Yoon, 2013). However, there was significant heterogeneity in effect sizes among included studies. The advantage in mental rotation for males was larger with shorter time limits on the mental rotation task.

Another cognitive domain that has been proposed to differ by sex/gender is verbal ability, where females are thought to have an advantage over males. A critical review of the literature found that although there is a slight advantage in language acquisition for young girls during development, this advantage appears to diminish in adulthood and there is otherwise no robust evidence for significant sex differences in verbal ability when large samples are used (Wallentin, 2009). However, sex/gender differences in verbal abilities continue to be investigated and contested. A more recent cross-sectional study found a female advantage in verbal fluency in children as well as adults (Barel & Tzischinsky, 2018). Further, a female advantage was also found in a sample of cognitively normal older adults without amyloid pathology on tasks of verbal memory, category fluency, and naming (Zhang et al., 2017).

With respect to memory, a large meta-analysis of 617 studies on sex/gender differences in episodic memory found a female advantage for episodic memory tasks involving verbal and non-verbal stimuli, while there was a male advantage for an episodic memory task involving spatial processing (Asperholm et al., 2019).

Overall, investigations of sex/gender and cognition reveal more similarities than differences. There have been several cross-sectional and longitudinal research studies that have been conducted to better understand sex/gender differences in cognitive aging.

What are the Sex and Gender Differences in Cognitive Aging?

Although sex and gender differences on specific cognitive abilities in healthy individuals have been small, questions still remain about sex/gender differences in cognitive changes that occur as part of healthy aging. Longitudinal study designs pose several advantages over cross-sectional studies as they can measure intraindividual change over time. A summary of the major longitudinal studies in healthy cognitive aging in older adults conducted thus far is provided in Table 1.3. Overall, most longitudinal studies have found minimal sex/gender differences in the rate of cognitive changes over time in older healthy adults.

A meta-analysis of 13 longitudinal studies concluded that sex/gender was not a determining factor in cognitive decline amongst individuals aged 60-80 (Ferreira et al., 2014). Although the meta-analysis did not find sex/gender differences in overall cognitive level, there was some consensus with respect to general sex/gender differences on specific cognitive tests that is consistent with previous findings. Females outperformed males on tests of episodic memory (de Frias et al., 2006; Wetherell et al., 2002) and males outperformed females on tests of visuospatial ability (de Frias et al., 2006; Lovden et al., 2004; Wetherell et al., 2002). Sex/gender differences for tests of semantic memory and attention were more inconsistent (Ferreira et al., 2014). The overall evidence suggests that there are few if any sex/gender differences in cognitive aging in healthy older adults before age 80. Other factors, like age, education, and social behaviour were more important in predicting cognitive changes than sex/gender.

However, a more recent study using data from the Baltimore Longitudinal Study on Aging found steeper rates of decline for males on measures of mental status, perceptuomotor speed and integration, and visuospatial ability, suggesting that females have higher resilience in age-related cognitive decline compared to males (McCarrey et al., 2016). Females showed higher baseline scores on tests of verbal learning and memory, fluent language production, mental status, and psychomotor speed, while males showed an advantage on tests of visuospatial ability, which could have influenced the findings. Both males and females showed a decline on all cognitive measures over the study period of 3-9 years.

Table 1.3 *Longitudinal Studies of Cognitive Aging in Healthy Older Adults*

Reference	Study	Participants	Duration	Findings	Sex/Gender Differences?
(Ferreira et al., 2014)	Systematic review	13 studies	5-17 years	Women > men in episodic memory Men > women in visuospatial ability	No
(Seeman et al., 2001)	MacArthur Studies of Successful Aging	Men=286 (73.8 ± 2.7 years) Women=420 (74.4 ± 2.7 years)	7.4 years ± 4.7 months	↓all aspects of cognitive performance	No
(Aartsen et al., 2004)	Longitudinal Study of Aging Amsterdam	Men = 1132 Women= 1175 (at least 60 years old at baseline)	6 years	women>men for immediate and delayed recall no sex/gender differences on coding	No
(Rabbitt et al., 2004)	Community sample from Newcastle-upon-Tyne and Greater Manchester	Men = 1735 Women = 4176 (age 49+; 76% >60 years)	17 years	↓AH4-1, AH4-2, CVL (accelerated with age) no change in MHA & MHB men > women for AH4-1, AH4-2, MHA, MHB women > men on CVL (linked with age) initial higher ability level associated with more rapid decline	No
(de Frias et al., 2006)	Betula Project	Men = 285 Women = 340 (52.8 ± 12.4 years)	10 years	Women > men on episodic recall, face and verbal recognition, and semantic fluency men > women on visuospatial ability	No
(Wetherell et al., 2002)	Swedish Adoption/Twin Study of Aging	Men = 173 (62.8 ± 8.0 years) Women=420 (64.4 ± 9.0 years)	6 years	women > men at figure identification, names and faces, and Thurstone's picture memory men > women at information, figure logic and card rotation	No
(Proust-Lima et al., 2008)	Personnes Agées Quid (PAQUID)	Men = 964 Women = 1264 (73.6 + 6.0 years)	13 years	No sex/gender differences in global cognition at age 65 women > men on verbal abilities men > women on visuospatial abilities	Women had slightly steeper declines in global cognition
(Singer et al., 2003)	Berlin Aging Study (BASE)	n = 132 (78.2 + 5.9 years) 55% women	6 years	Perceptual speed, memory, and fluency showed similar amounts of decline, whereas knowledge showed preserved stability until about age 90	No
(Lovden et al., 2004)	Community sample from Sweden	Men = 171 Women = 190 Age range 60-80 years	5 years	greater episodic and semantic memory decline with advancing age higher education associated with better memory performance women > men for episodic memory	No

(McDowell et al., 2004)	Canadian Study of Health and Aging	Analysis I: n = 3,255 (M ~ 73 years) [~ 63% women] Analysis II: Community n = 5,879 (M = 74.3 years) [60.9% women] Institution n = 223 (M = 78.5 years) [75.3 % women]	Analysis I: 10 years Analysis II: 5 years	almost 50% of elderly participants showed virtually no change in global cognition over 10 years	No
(Barnes et al., 2003)	Religious Orders Study	Men = 271 (74.0 + 6.5 years) Women = 577 (76.1 + 7.0 years)	8 years	↓all tests of cognitive performance over 8-year period risk of AD did not differ between men and women	No
(Finkel et al., 2003)	Swedish Adoption/Twin Study of Aging	n=590 (65.1 ± 8.0 years) [60.7% women]	6 years	At age 65, men > women for general knowledge, spatial reasoning, and spatial ability	No
(Finkel et al., 2006)	Swedish Adoption/Twin Study of Aging	n = 798 (Age 44-88 at 1st testing wave) 60% women	13 years	women > men on processing speed and memory men > women on information, spatial ability	Faster decline for women on information; faster decline for men on card rotation; no consistent indications that sex/gender patterns of cognitive aging
(Gerstorf et al., 2006)	Berlin Aging Study (BASE)	n = 368 (mean age = 83 years)	13 years	↓all tests of cognitive performance over study period women > men on Digit Letter, Paired Associates, Memory for Text, Word Beginnings, and Spot-a-Word	No
(McCarrey et al., 2016)	Baltimore Longitudinal Study of Aging	n = 1,065 to 2,127 (mean age = 64.1 to 69.7 at baseline)	Mean follow-up = 3-9 years	↓all tests of cognitive performance over study period women > men on verbal learning and memory, fluent language production tests of mental status, psychomotor speed and perceptual integration men > women on visuospatial ability	Faster decline in men for mental status, perceptual integration and speed, and spatial ability

Research Aims and Questions

The overall goal of this research is to better understand structural brain changes and changes in cognitive functioning in healthy adults and how these changes may vary by biological sex. This manuscript-based dissertation contains three complementary studies to answer the following research questions:

1) What are the sex differences in white matter microstructure among healthy older adults?
2) How does white matter microstructure change over time among healthy older adults and are there differences in rates of change for males and females?
3) What are the sex similarities and differences in cognition and rates of cognitive change over time among healthy adults?

Common Data Source

All three studies involved secondary data analyses of neuroimaging and neuropsychological data that were collected and made available to researchers by the Parkinson Progression Markers Initiative (Sullivan et al., 2001). The PPMI is a comprehensive observational, international, multi-center study designed to identify progression biomarkers for Parkinson’s disease (http://www.ppmi-info.org/). As part of this initiative, the open-access database provides clinical, molecular, and imaging data for a variety of Parkinson’s disease cohorts. Although the primary focus of the PPMI is to investigate Parkinson's disease, the study also collected data on healthy individuals as part of its normal control group. This provided a convenient opportunity to investigate structural and functional changes among healthy adults. Participants for the present research were selected from a cohort of healthy adults from the PPMI control group, who were healthy adults over the age of 30 without a first-degree relative with Parkinson’s disease. PPMI's complete initial inclusion and exclusion criteria can be found in **Appendix A**. Participants from the PPMI control group were excluded from all three studies in the following cases: a) if PPMI’s criteria for test-based mild cognitive impairment were met (i.e., two or more cognitive test scores > 1.5 standard deviations below the mean), or b) if there was a PPMI investigator diagnosis of mild cognitive impairment or dementia. Additional exclusion criteria will be discussed for each individual study.

The PPMI Operations Manual indicated that they collected a sex/gender variable based on “gender at birth”. Based on the consensus definitions outlined in Table 1.1, it is apparent that PPMI’s use of “gender” corresponds with the consensus definitions of “sex” and this variable was treated as such. Unfortunately, PPMI did not collect data corresponding to the consensus definitions of “gender”, so it was not possible to examine the effects of gender within the database.

Ethical approval was obtained by the PPMI study, and ethical approval from the University of Victoria Ethics Board for secondary data analysis was obtained. All manuscripts resulting from this research will include the following PPMI Publication Statement: Data used in the preparation of this article were obtained from the Parkinson’s Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson’s Research and funding partners, which can be found at www.ppmi-info.org/fundingpartners.

Overview of the Studies

Study 1 and 2 utilized neuroimaging data (i.e., DTI data) to compare white matter microstructure between males and females and changes in white matter microstructure over time. Study 3 focused solely on the neuropsychological data to investigate longitudinal changes in cognition and sex differences in rates of cognitive change over time. A brief overview of these studies is provided.

Study 1: Sex Differences in White Matter Microstructure in Healthy Older Adults

In Study 1, sex differences in white matter microstructure among a sample of healthy older adults were compared at a single time point. Specifically, tract-based spatial statistics were used to compare whole brain FA and MD values between older adult males and females with available DTI data from the healthy control group of the PPMI.

Study 2: Longitudinal Changes in White Matter Microstructure in Healthy Older Adults

In Study 2, changes in white matter microstructure were examined longitudinally over one year. Participants (N=39) included 16 females (mean age = 60.50 ± 5.99) and 23 males (mean age = 65.50±7.53) from the healthy control cohort of the PPMI. DTI data from two time points (baseline and approximately one year later) were analyzed using tract-based spatial statistics to examine longitudinal changes in FA and MD.

Study 3: Sex Similarities and Differences in Cognition: A Longitudinal Study of Healthy Control Participants from the Parkinson's Progression Markers Initiative

In Study 3, longitudinal data from the healthy control group of the PPMI was used to examine cognitive changes in healthy adult men (n=125, mean age=61.61, SD=10.97) and women (n=68, mean age=59.44, SD=11.56). Participants completed neuropsychological measures annually for up to five years (i.e., six time points). These measures included the Montreal Cognitive Assessment, Letter Number Sequencing, Semantic Fluency, Symbol Digit Modalities Test, Benton Judgment of Line Orientation, Hopkins Verbal Learning Test-Revised Immediate and Delayed Recall. Neuropsychological measures were modeled separately as a single outcome variable in a series of 2-level linear mixed models. Within-person changes in cognition over time (i.e., level-1) and between-group differences in longitudinal change trajectories as predicted by sex (i.e., level 2) were examined. The effects of age and education were also investigated.

These studies will contribute to a growing body of research investigating changes in the structure of the brain and in cognition in the context of healthy aging. Further, these studies will explicitly contribute to our knowledge of sex and gender differences in healthy aging with the explicit examination of sex similarities and differences in white matter microstructure and cognition both at baseline and over time. Findings will aid in the promotion of healthy and successful aging throughout adulthood and into the latter decades of life.

Chapter 2: Study 1 - Sex Differences in White Matter Microstructure in Healthy Older Adults

**This chapter is formatted to submit for publication.*

Abstract

Background: Despite white matter comprising 50% of brain tissue, there has been limited research on sex differences in white matter microstructure in healthy older adults. Further, given that males tend to have larger brain volumes than females, some studies may have yielded potentially spurious findings by failing to adjust for total brain volume (TBV). This study aimed to examine sex differences in white matter microstructure in healthy older adults while controlling for TBV.

Methods: Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). Diffusion weighted images for 16 female and 24 male healthy cohort participants were compared using tract-based spatial statistics on two of the most frequently used diffusion tensor imaging metrics: fractional anisotropy (FA) and mean diffusivity (MD).

Results: Males were significantly older and had significantly larger TBVs than females. There were no significant differences in whole brain FA or MD between males and females when controlling for age and TBV.

Conclusions: No sex differences in white matter microstructure were found in a sample of healthy older adults. These findings add to a growing body of literature that suggest that there are limited sex differences in brain structure in healthy individuals. Future research should investigate sex-specific changes in white matter microstructure longitudinally, across the healthy aging trajectory.

Keywords: diffusion tensor imaging, sex, healthy aging, white matter microstructure,

Introduction

Sex and gender are important determinants of health and aging (Banister et al., 2012; Barford et al., 2006). Women outlive men in virtually all populations globally, with an average global life expectancy of 74.2 years for women and 69.8 years for men (World Health Organization, 2020). There are also sex and gender differences in the prevalence of many disorders of the central nervous system. For example, Alzheimer's disease and Multiple Sclerosis are more prevalent in women, while Parkinson's disease and Amyotrophic Lateral Sclerosis are more prevalent in men (Zagni et al., 2016). Neuroimaging affords non-invasive methods to examine the structure of the brain in vivo. Better understanding possible sex and gender differences in brain structure in healthy individuals could assist in the promotion of healthy aging and prevention of age-related diseases.

A major challenge for studying sex and gender differences is the inconsistent use of terminology throughout the literature. However, over the last decade, several national and international organizations have published consensus definitions of sex and gender. These definitions agree that *sex* refers to biological and physical characteristics of humans, whereas *gender* refers to socially constructed characteristics of personal identity (Government of Canada, 2014; National Institutes of Health, n.d.; World Health Organization, n.d.). Others have argued that the composite term *sex/gender* should be used to the reflect the difficulty, if not impossibility, of separating inborn sex factors from social and environmental gender learning (Kaiser, 2012; Rippon et al., 2014). Given the inconsistency in terminology throughout the literature, we will use the terms sex and gender when one is clearly identified as the variable of interest with a stated meaning that agrees with the consensus definitions; the term *sex/gender* will be used when the distinction is unclear.

When considering brain structure and function, evidence shows that males and females are more similar than they are different. A comprehensive synthesis of the literature concluded that the sex/gender differences in brain structure and function are very small after accounting for differences in overall brain volume (Eliot et al., 2021). On average, males are taller, weigh more, and have larger heads than females. This trend continues neuroanatomically, where males have larger brain sizes on average than females. This is most often measured as total intracranial volume (ICV), which includes all the anatomy within the skull, and total brain volume (TBV), which includes all the anatomy within the skull excluding the meninges and ventricles. Males have been found to have 11% larger ICVs and 12% larger TBVs than females (Ruigrok et al., 2014). Males have also been found to have significantly larger volumes in many respective parts of the brain compared to females, including 10% larger cerebrums, 9% more gray matter, 13% more white matter, 11.5% more cerebrospinal fluid, and 9% larger cerebellums. Overall, there appears to be minimal sex/gender differences in brain structure after accounting for larger brain volumes in males (Eliot et al., 2021). Contrary to these results, a more robust analysis of over 40,000 participants from the UK Biobank and concluded that sex/gender differences in height accounted for approximately 39% of sex/gender differences in TBV and 409/620 brain measures showed small, but significant sex/gender differences, even after adjusting size for the corresponding global measure of TBV, total surface area, or total mean cortical thickness (Williams et al., 2021). Although most of the structural brain differences were small (i.e., mean effect size of d=0.13), it is important to continue to explore these sex/gender differences and their corresponding effects on health and aging.

White matter is comprised of myelin sheathed neuronal axons that connect structures throughout the brain. Despite representing approximately 50% of the brain, studies of sex/gender differences in white matter microstructure have been limited and are worthy of further investigation. Diffusion tensor imaging (DTI) is a magnetic resonance imaging (MRI) technique that uses the principals of water diffusion to infer the integrity of white matter microstructure throughout the brain. The mostly commonly used DTI metrics include fractional anisotropy (FA), which represents the overall directional restriction of diffusion of water, and mean diffusivity (MD), which represents the overall diffusivity within tissue (Soares et al., 2013). In general, decreased FA reflects decreased anisotropy or directionality of diffusion and increased MD represents greater overall diffusion, both of which suggest less restriction of diffusion and decreased white matter microstructural integrity.

DTI studies investigating sex/gender differences in white matter microstructure are limited and have mixed findings. Several studies comparing males and females have found inconsistent sex/gender differences in DTI metrics throughout various regional structures in the brain (Chou et al., 2011; Cox et al., 2016; Kanaan et al., 2012, 2014; Menzler et al., 2011). For example, in an investigation of white matter differences in 80 healthy adults from Taiwan, Chou et al., (2011) found females had greater FA in the fronto-occipital fasciculus, body of the corpus callosum, and white matter underlying the parahippocampal gyrus, while males had greater FA in the bilateral internal capsule, white matter underlying the medial frontal gyrus, fusiform gyrus, hippocampus, insula, postcentral gyrus, frontal and temporal lobe. Kanaan and colleagues conducted two studies (2012, 2014) in which males were found to have higher FA in the superior cerebellar peduncles and females were found to have higher FA in the corpus callosum. Menzler et al. (2011) also compared sex/gender differences in white matter microstructure in a sample of 50 healthy adults, finding increased FA and decreased radial diffusivity in males in the thalamus, cingulum, and the corpus callosum. However, these studies were conducted with healthy adults in young to middle adulthood (i.e., 20-40 years) and thus the results cannot be assumed to be the same for older populations. One of the largest samples to date comes from a UK Biobank Study, which included 3513 healthy older adults from age 44.64 to 77.12 years (Cox et al., 2016). Significant sex/gender differences were found in overall FA (males>females) and neurite orientation dispersion (a measure of spatial configuration of neurite structures) (females>males), and several other metrics in various directions in 15 specific regions of interest. A major issue that could explain these inconsistent findings is the lack of consensus on whether to control for overall brain size when conducting sex/gender-based analyses of brain structure (Kanaan et al., 2012) and the appropriate methodology to do so. Many DTI studies do not appear to adequately address this issue. However, some have found that most of the significant sex/gender differences in white matter microstructure diminished after adjusting for total ICV (Takao et al., 2014), while others have continued to find some sex/gender differences in portions of large white matter structures, such as the corpus callosum, even when accounting for sex/gender differences in overall callosal size (Westerhausen et al., 2011). Taken together, no consistent sex/gender differences in white matter microstructure have been observed. Inconsistent definitions of sex and gender, varying sample size and characteristics (e.g., age), as well as varying methodological approaches (e.g., imaging parameters, statistical methods) make comparison across studies difficult. Thus, questions remain as to whether healthy older adult males and females differ in overall white matter microstructure once adequately controlling for gross morphological differences in brain size.

The present study aimed to investigate sex differences in white matter microstructure in healthy older adults. Specifically, we aimed to answer whether there are differences in FA and MD between males and females as measured by DTI. In accordance with the literature that finds more similarities than differences in brain structure between the sexes after controlling for TBV, it was predicted that there would be limited differences in FA or MD between males and females after adjusting for TBV.

Method

Participants

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (Marek et al., 2011b). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, which can be found at www.ppmi-info.org/fundingpartners**.** Participants were eligible for inclusion in the present study if they were in the healthy cohort, with available diffusion weighted imaging data. Participants were excluded from the analysis if: a) PPMI's criteria for test-based mild cognitive impairment were met (i.e., two or more cognitive test scores were > 1.5 standard deviations below the mean), b) there was a PPMI investigator diagnosis of mild cognitive impairment or dementia, c) the imaging data from was too poor in quality to be included in the analysis, as determined by visual inspection. It is important to note that the PPMI Operations Manual indicated that they collected a sex/gender variable based on "gender at birth". Based on the consensus definitions in the literature, it is apparent that PPMI's use of "gender" corresponds with the consensus definitions of "sex" and this variable was treated as such. Unfortunately, PPMI did not collect data corresponding to the consensus definitions of "gender", so it was not possible to examine the effects of gender within this database. Ethical approval was obtained by the PPMI study, and the current study also obtained ethical approval from the University of Victoria Ethics Board for secondary data analysis. These criteria yielded a total of 40 participants (16 female and 24 male). Demographic information is available in Table 2.1.

Diffusion-Weighted Images

All images were acquired with a Siemens 3T TIM Trio scanner with a 12 channel Matrix head coil. Diffusion-weighted images were acquired with a spin echo, echo planar imaging sequence with 64 gradient directions and a b-value of 1000s/mm^2 with a voxel size of 2 mm^3. For more information regarding MRI acquisition, please see: http://ppmi-info.org/.

Procedure and Data Analysis

Diffusion-weighted imaging data and demographic information of eligible participants were downloaded from the PPMI website (www.ppmi-info.org/data). Descriptive statistics were conducted using R-Studio Version 1.1.419 and fMRI Software Library (FSL) was used for all image preprocessing and imaging analyses (Jenkinson et al., 2012; Smith et al., 2004a; Woolrich et al., 2009).

Image Preprocessing

Raw diffusion-weighted images were converted from DICOM to NifTi format using dcm2nii converter from mricron (Rorden & Brett, 2000). First, to correct for eddy currents distortions and head movement, the Eddy Current Correction tool (J. L. R. Andersson & Sotiropoulos, 2016) was utilized. Next, the skull and other non-brain tissue was removed from the images using the Brain Extraction Tool (Smith, 2002) with visual inspection for accuracy.

Tract-based Spatial Statistics

Tract-Based Spatial Statistics (TBSS) (S. M. Smith et al., 2006), within FSL (Smith et al., 2004a), is a fully automated approach to objectively estimate the overall white matter tracts within the brain that are common to study subjects, which can then be compared statistically. First, FA images were created by fitting a tensor model to the raw diffusion data using DTIfit (Behrens et al., 2003; Johansen-Berg et al., 2004). All subjects' FA data was aligned into a common space using the nonlinear registration tool FNIRT (Andersson et al., 2007), which uses a b-spline representation of the registration warp field (Rueckert et al., 1999). Next, the mean FA image was created and thinned using a threshold of 0.2 to create the mean FA skeleton, which represents the centers of all tracts common to the group. Each subjects’ aligned FA data were then projected onto the mean FA skeleton and the resulting data were used for voxelwise cross-subject statistics. The previous steps were repeated for MD.

Voxelwise Statistical Comparisons

Analyses were conducted to examine between-group differences in FA and MD at the whole brain level. All contrast files were created using FSL’s General Linear Model. A two-group difference model was created and adjusted for two covariates: age and TBV. Age and TBV were mean centered by subtracting the overall group mean age and TBV from each individual age and TBV, respectively. TBV values were calculated by segmenting T1 MRI images using FSL’s Automated Segmentation Tool (FAST) and adding the resultant volumes of grey matter and white matter for each participant (Zhang et al., 2001). Voxelwise comparisons were conducted using FSL’s Randomise, which is a tool for nonparametric permutation inference on neuroimaging data (Winkler et al., 2014). Randomise was conducted using threshold-free cluster enhancement to correct for multiple comparisons, with the recommended 5000 permutations (Smith & Nichols, 2009). Randomise produced test static images for each contrast of interest (i.e., females>males and females<males), which were individually examined by overlaying the respective image on its corresponding mean skeleton mask and standard brain image within FSLeyes, the FSL image viewer. This process was conducted for both FA and MD.

Results

Descriptive statistics regarding age, education, TBV, and whole brain FA and MD values can be found in Table 2.1. Males were significantly older than females and had significantly larger TBVs. Results of voxelwise comparisons revealed no statistically significant differences in FA or MD between males and females when controlling for age and TBV.

Table 2.1 *Participant Demographics*

Variable	Females (n=16)	Males (n=24)	*df*	*t*	*p*	Cohen's d
Mean Age (±SD)	**60.51±5.96**	**65.50±7.53**	**36.78**	**-2.33**	**<*.05***	**0.72**
Min-Max Age	50.27-70.18	52.10-79.83				
Mean Education (±SD)	15.00+3.14	16.08+2.41	26.45	-1.17	0.25	0.4
Mean TBV (±SD)	**1012880 ±62723.11**	**1137737 ±85366.96**	**37.56**	**-5.33**	**<*.05***	**1.62**
Mean FA (±SD)	0.46±0.02	0.46±0.02	36.927	-0.68	0.49	0.21
Mean MD (±SD)	0.000716	0.000716	36.06	0.02	0.98	0.0069
Race/Ethnicity (number)						
White	15	23				
Black	0	1				
Asian	0	0				
Other	1	0				

Note. TBV=total brain volume, FA=fractional anisotropy (no units, values range from 0, reflecting complete isotropy to 1, reflecting complete anisotropy), MD=mean diffusivity (mm^2/sec). Mean FA and MD values reflect the average of individual mean FA or MD values within all the white matter tracts in the brain. These values were unadjusted for age and TBV.

Discussion

This study aimed to determine if there were sex differences in white matter microstructure in healthy older adults. In this sample of 16 female and 24 males selected from the healthy control group of the PPMI database, males were significantly older and had significantly larger TBVs than females. The results of the TBSS analyses revealed no detectible sex differences in FA or MD when controlling for age and TBV. These results were congruent with our hypotheses and agree with existing literature that finds more similarities than differences in brain structure between the sexes after controlling for overall larger brain volume of males (Eliot et al., 2021).

There has been much interest in studying sex differences in brain structure throughout the lifespan. This is particularly important throughout the latter decades of life given the significant disparities between males and females in age-related diseases and overall life expectancy. Although white matter represents approximately 50% of brain tissue, studies investigating sex/gender differences in its microstructural integrity have been limited and findings have been mixed. Several DTI studies have found significant differences between males and females throughout various structures in the brain (Chou et al., 2011; Kanaan et al., 2012, 2014; Menzler et al., 2011). While these studies also utilized TBSS, comparison to the present results is difficult as the samples included participants who were at least two decades younger on average. A robust study of aging and white matter using a sample of similarly aged older adults from the UK Biobank controlled for head size and found overall higher FA in males than females, overall higher neurite orientation dispersion in females, and overall similar MD between males and females (Cox et al., 2016). There were also small but significant differences on several DTI metrics in various directions throughout 15 specific regions of interest. It is important to note that these differences, although significant, were found to be small. Large samples sizes are likely needed to detect small differences, and this could explain the variability in findings among the literature. Variability in acquisition parameters, neuroimaging software, and analytical approaches could also contribute to mixed findings.

We found no significant sex differences in whole brain FA or MD. While others have continued to find some small sex/gender differences in portions of specific large white matter structures, such as the corpus callosum, even when accounting for sex/gender differences in overall callosal size (Westerhausen et al., 2011), others have also found that sex/gender difference in white matter microstructure diminish to non-significance when controlling for similar measurements of brain size, such as ICV (Takao et al., 2014). It is difficult to estimate the number of unpublished DTI studies that may have also found no significant sex differences in white matter microstructure due to the “file drawer problem”. Indeed, there is evidence of potential reporting bias towards “positive” (i.e., significant) sex/gender differences in functional magnetic resonance imaging studies (David et al., 2018). Although to our knowledge no studies have been conducted to estimate the specific reporting biases for DTI studies, it is very likely that full studies or portions of studies that include sex/gender-based analyses of white matter were not submitted for publication due to null findings or were rejected during the review process.

Consistent with the literature (Eliot et al., 2021; Williams et al., 2021), males were found to have significantly larger TBVs than females in the present study. TBV was included in our models as a covariate based on evidence that FA and MD values are affected by fiber bundle thickness, orientation, and curvature that can modulate partial volume effects that may mask true microstructural group differences (Vos et al., 2012). While some researchers have debated whether these differences should be considered a confounding or explanatory factors (Kanaan et al., 2012), most do not discuss or address the issue at all, and there is no consensus among researchers of how to account for the consistent sex/gender difference in overall brain volume. Males generally have greater mass than females and thus likely require a greater volume of corresponding neuronal tissue. It is important to understand what brain differences remain after controlling for these gross anatomical differences between males and females. It is possible that previous significant findings of sex differences may be explained or partially explained by larger brain volume in males.

The current study had several limitations. First, the sample was limited to participants from the PPMI database with available diffusion weighted imaging data, which limited its size and racial/ethnic diversity. Second, it was only possible to examine sex differences (i.e., male or female biological sex assigned at birth) in white matter microstructure as information about gender was not collected. Some researchers have built on these recommendations by constructing gender-based metrics using proxy variables to predict health events using large databases (Levinsson et al., 2022). Using this approach, it may be possible to retrospectively create gender-based variables in the PPMI and other large neuroimaging databases, which may lack explicit gender-based variables. Going forward, future research should follow new recommendations to ensure comprehensive and clear reporting of sex and gender in clinical research (Clayton & Tannenbaum, 2016).

Despite these limitations, we found no significant sex differences in whole brain FA or MD in a sample of healthy older adults. Age and TBV were included as covariates in our models as males were significantly older and had larger TBVs than females. Limited research has investigated sex differences in older healthy individuals and these results should be replicated in larger samples with appropriate adjustments for overall brain size. Future research should continue to investigate white matter differences among healthy individuals of varying sex and gender, especially among individuals who are transgender, who have often been underrepresented in research. Additional research may also examine how white matter microstructure may change throughout the normal aging trajectory and whether there are sex or gender specific rates of change.

Chapter 3: Study 2 - Longitudinal Changes in White Matter Microstructure in Healthy Older Adults

Abstract

Background: As the global population of older adults increases, it is crucial to study the healthy aging brain. Despite white matter (WM) comprising approximately 50% of brain tissue, limited research has investigated the expected changes that occur during the normal aging process in healthy older adults. Longitudinal studies pose several advantages over cross-sectional methods and allow the examination of intraindividual change over time. The Parkinson's Progression Markers Initiative (PPMI) is an ongoing observational longitudinal study designed to investigate biomarkers related to Parkinson's disease. These data present a convenient opportunity to investigate normal brain changes among healthy older adults by using data collected on the healthy control cohort. The primary aim of the current study was to investigate longitudinal changes in WM using diffusion tensor imaging. Determining whether there were sex differences in WM change over time was a secondary objective of the study.

Methods: Participants (N=39) included 16 females (mean age = 60.50 ± 5.99) and 23 males (mean age =65.50±7.53) from the healthy control cohort of the PPMI. DTI data from two time points (baseline and approximately one year later) were analyzed using tract-based spatial statistics from the FMRIB Software Library (FSL). Diffusion weighted images were acquired with a Siemens 3T TIM Trio scanner with a 12 channel Matrix head coil. All images were acquired with a spin echo, echo planar imaging sequence with 64 gradient directions and a b-value of 1000s/mm^2 with a voxel size of 2 mm^3. A within-subjects design was employed to examine longitudinal changes in WM from baseline to one year later for all participants, males only, and then females only. DTI metrics included fractional anisotropy (FA) and mean diffusivity (MD).

Result: Longitudinally over one year, decreases in MD ($p<.05$, corrected) were found in the right superior and inferior longitudinal fasciculus, the right corticospinal tract, and the right inferior fronto-occipital fasciculus. Stability in FA was observed over one year. When examining changes in WM separately by sex, stability in FA and MD was found for both groups. Individual trajectories of FA and MD change were also examined.

Conclusion: Changes in WM can be captured via neuroimaging across a one-year time frame in healthy older adults. Post hoc examination of the current results suggests great intraindividual variability in trajectories over time, which may have affected the overall group trajectory. Continued research of factors that contribute to the identifying individual healthy aging trajectories is warranted.

Keywords: healthy aging, diffusion tensor imaging, mean diffusivity, fractional anisotropy, neuroimaging, sex differences

Introduction

The population over the age of 65 is expected to double from 703 million to 1.5 billion by 2050, and the population of adults over the age of 80 is expected to triple from 143 million to 426 million by 2050 (United Nations et al., 2019). Notably, age is a primary risk factor for many types of dementia (Hou et al., 2019), including Alzheimer's disease (Reitz & Mayeux, 2014), Lewy Body Dementia (Jellinger, 2003), and Parkinson's disease (Reeve et al., 2014). The number of individuals with dementia is also expected to rise significantly with the growing aging population. Given this, it is even more vital to understand changes in the aging brain to inform and guide the development and enhancement of interventions that promote health and longevity into the later decades of life.

A great deal of research has focused on disease specific changes in the brain. However, the expected structural changes that occur as part of normal healthy aging are not yet entirely understood, especially with respect to white matter and as a function of sex. White matter comprises the neuronal axons that support neural connectivity throughout the brain. Although white matter comprises approximately 50% of brain tissue, studies of changes in white matter microstructure throughout the normal aging process have been limited. Diffusion tensor imaging (DTI) is a magnetic resonance imaging (MRI) technique, which uses diffusion weighted imaging to quantify the molecular diffusion of water molecules to infer in vivo information about white matter microstructural integrity. The most commonly used DTI metrics include fractional anisotropy (FA), which provides an estimate of the degree of anisotropic diffusion, and mean diffusivity (MD), which describes the average diffusion within a voxel. Other metrics include axial diffusivity (AD) and radial diffusivity (RD), which measure diffusivity across the principal axis and the average of the two minor axes, respectively.

Recent efforts have been made to quantify normal expected age-related structural brain changes, including white matter, across the lifespan into normative brain charts (Bethlehem et al., 2022). Findings aggregated from over 100 cross-sectional primary studies and 101,457 participants from 115 days post-conception to 100 years of age showed that white matter volume increased rapidly from mid-gestation to early childhood, peaked at approximately age 28.7 years, before an accelerated decline after age 50. Similarly to the normative brain charts (Bethlehem et al., 2022), a systematic review of 30 DTI largely cross-sectional, but also longitudinal studies has summarized the known changes in specific white matter microstructural metrics in healthy individuals across the lifespan: white matter development shows a curvilinear pattern where there is a rapid increase in FA during the neonatal period, which continues into childhood and adolescence, before slowing and plateauing in mid-adulthood, then declining with old age (Yap et al., 2013). This has been corroborated by others who have found that FA peaks in most tracts between the ages of 23 to 39 (Kochunov et al., 2012). In general, MD has shown the opposite trend of FA, where MD decreases during the neonatal period, and continues to decline in childhood, adolescence and young adulthood, before increasing in older adults and the elderly(Yap et al., 2013). Regionally, it appears that peak FA is reached in associative white matter tracts in later life and shows more age-related decline than areas that reach peak FA at a younger age, such as the motor and sensory tracts (Kochunov et al., 2012; Yap et al., 2013). A “last-in-first-out” hypothesis of brain aging has emerged, which posits that the brain areas to develop last are the most vulnerable to degeneration (Raz et al., 2007; Raz & Rodrigue, 2006).

Several longitudinal studies examining change in white matter microstructure over time have been conducted. An overall pattern of annual decreases in white matter microstructure emerged in a longitudinal study of 203 adults from age 20-84 years, where change in DTI metrics were examined annually over 3.6 years (Sexton et al., 2014). Older age was associated with greater changes in widespread regions, including the occipital and temporal lobes, and more prominently the frontal and parietal lobes. Findings from several other longitudinal studies also show that there are widespread changes in white matter microstructure with age. Most prominently observed were decreases in FA (Bender et al., 2016; Charlton et al., 2010; Lövdén et al., 2014a; Sexton et al., 2014) and increases in MD (Charlton et al., 2010; Lövdén et al., 2014b; Sexton et al., 2014).

Regarding sex differences in brain structure, many studies have found sex differences in certain brain regions at single time points, but a review of the literature suggest that many of these differences diminish after controlling for overall brain size (Eliot et al., 2021). Across the lifespan, males have been found to have larger white matter volumes than females at most stages of life after early childhood, but this was not thought to be indicative of any difference in clinical or cognitive outcomes (Bethlehem et al., 2022). Further questions remain as to whether males and females age in similar ways. That is, do males and females have similar changes in brain structure over time? Sex differences in the prevalence and course of many age-related diseases exist, which suggests that there may also be sex differences in the healthy aging process. Limited research has examined sex specific differences in the rates of white matter change over time in healthy adults.

Cross-sectionally, middle-aged adults have been found to have lower FA and AD and higher MD and RD than younger adults; however, these changes did not differ by sex (Ouyang et al., 2021). Longitudinally, one of the largest samples to date comes from a UK Biobank Study (Cox et al., 2016), which included 3513 healthy older adults from age 44. 64 to 77.12 years. With respect to white matter microstructure, older age was associated with decreased microstructural integrity in the thalamic radiations and association fasciculi. MD was found to be the most age-sensitive measure parameter. Regarding sex, males were found to have higher FA than females, while females were found to have overall higher neurite orientation dispersion (a measure of spatial configuration of neurite structures), than males. Small and inconsistent interactions between age and sex were found. Overall, limited research has explicitly examined sex differences in white matter microstructural change over time. Only about 10% of the studies used to create normative brain charts, were longitudinal (Bethlehem et al., 2022), which highlights a significant gap in the literature.

Longitudinal study designs are necessary for measuring intraindividual change over time. Longitudinal DTI data from healthy older adults was used to investigate the expected aging trajectory in white matter microstructure among healthy older adults. There were two primary research questions:

1) How does white matter (i.e., FA and MD) change in healthy older adults over one year?
2) Are there sex differences in changes in white matter (i.e., FA and MD) in healthy older adults over one year?

It was hypothesized that decreases in FA and increases in MD would be observed, corresponding to the discussed literature. It is also possible that stability in these metrics would be found due to the short follow up time frame. Regarding longitudinal changes by sex, it was hypothesized that males and females will show similar stability or decline over time, based on the limited research available. The results from this study will add to a growing body of literature documenting the expected age-related trajectory of structural brain changes in healthy older adults, and whether these age-related trajectories differ between males and females.

Methods

Participants

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, which can be found at www.ppmi-info.org/fundingpartners. Participants were eligible for inclusion in the present study if: a) they were a healthy control group participant, and b) diffusion-weighted imaging data was available for the baseline and one-year follow-up date. PPMI's criteria for the healthy control group can be found in Appendix A. Participants were excluded from the analysis if: a) PPMI's criteria for test-based mild cognitive impairment were met (i.e., two or more cognitive test scores > 1.5 standard deviations below the mean), b) there was a PPMI investigator diagnosis of mild cognitive impairment or dementia, c) the imaging data from either time point was too poor in quality to be included in the analysis, as determined by visual inspection. Ethical approval was obtained by the PPMI study, and the current study has obtained ethical approval from the University of Victoria Ethics Board for secondary data analysis. These criteria yielded a total of 40 participants (16 female and 24 male). Demographic information is available in Table 3.1.

Diffusion-Weighted Images

All images were acquired with a Siemens 3T TIM Trio scanner with a 12 channel Matrix head coil. Diffusion-weighted images were acquired with a spin echo, echo planar imaging sequence with 64 gradient directions and a b-value of 1000s/mm^2 with a voxel size of 2 mm^3. For more information regarding 167 MRI acquisition, please see: http://ppmi-info.org/.

Procedure and Data Analysis

Diffusion-weighted imaging data for two timepoints (i.e., baseline and one-year follow-up) and demographic information of eligible participants were downloaded from the PPMI website (www.ppmi-info.org/data). Descriptive statistics were conducted using R-Studio Version 1.1.419 and image pre-processing and analyses were conducted using FMRIB Software Library (FSL) Version 5.0.10 tools (Smith et al., 2004b). DTI metric values were extracted from FSL and plotted using the ggplot2 package in R.

Image Preprocessing

Raw diffusion-weighted images were converted from DICOM to NifTi format using dcm2nii converter from mricron (Rorden, 2016; www.nitrc.org/projects/mricron). FMRI Software Library (FSL) Version 5.0.10 was used for all image preprocessing and imaging analyses (Smith et al., 2006). First, to correct for eddy currents distortions and head movement, the Eddy Current Correction (ECC) tool was utilized (J. L. R. Andersson & Sotiropoulos, 2016). Next, the skull and other non-brain tissue was removed from the images using the Brain Extraction Tool (BET) (Smith, 2002) with visual inspection for accuracy.

Tract-based Spatial Statistics

Tract-Based Spatial Statistics (TBSS), a tool within FSL (Smith et al., 2004a), is a fully automated approach to objectively estimate the overall white matter tracts within the brain that are common to study subjects, which can then be compared statistically. First, fractional anisotropy (FA) images were created by fitting a tensor model to the raw diffusion data using DTIfit (Behrens et al., 2003; Johansen-Berg et al., 2004). All subjects' FA data was aligned into a common space using the nonlinear registration tool FNIRT (Andersson et al., 2007), which uses a b-spline representation of the registration warp field (Rueckert et al., 1999). Next, the mean FA image was created and thinned using a threshold of 0.2 to create the mean FA skeleton, which represents the centers of all tracts common to the group. Each subjects' aligned FA data were then projected onto the mean FA skeleton and the resulting data were used for voxelwise within-subject statistics. The previous steps were repeated for MD**.**

Voxelwise Statistical Comparisons

Analyses were conducted to examine within-subject differences in FA and MD between the baseline and follow-up time point at the whole brain level for all participants, males only, and females only. All contrast files were created using FSL's General Linear Model. A repeated measures ANOVA model was created with a fixed factor of time with two levels (i.e., baseline and follow-up) and a random factor of subject. Each subject had a random intercept, which was modeled with dummy or indicator variables. The model was created such that there were two contrasts (i.e., C1 = baseline > follow-up, C2 = follow-up > baseline). Voxelwise comparisons were conducted using FSL's Randomise, which is a tool for nonparametric permutation inference on neuroimaging data (Winkler et al., 2014). Randomise was conducted using threshold-free cluster enhancement to correct for multiple comparisons, with the recommended 5000 permutations (Smith & Nichols, 2009). For each contrast, Randomise produced a test static image that was overlaid onto its corresponding mean skeleton mask and standard brain image within FSLeyes, the FSL image viewer., which were examined with visual inspection. There was a total of 12 contrast files, where there were two contrasts files (i.e., C1 = baseline > follow-up, C2 = follow-up > baseline) for each analysis comparing the baseline DTI metric to the DTI metric one-year later using all participants, males only, or females only. These analyses were conducted for both FA and MD. Specifically, the contrasts were:

1. Analyses for FA:

   All participants: C1 = baseline > follow-up, C2 = follow-up > baseline

   Males only: C1 = baseline > follow-up, C2 = follow-up > baseline

   Females only: C1 = baseline > follow-up, C2 = follow-up > baseline

2. Analyses for MD:

   All participants: C1 = baseline > follow-up, C2 = follow-up > baseline

   Males only: C1 = baseline > follow-up, C2 = follow-up > baseline

   Females only: C1 = baseline > follow-up, C2 = follow-up > baseline

Sex differences were investigated via visual inspection of contrast images with corresponding statistical thresholding at $p<.05$.

Results

Table 3.1 displays descriptive statistics for the study participants. At baseline, males were significantly older than females by approximately five years ($t(38)=-2.47$, $p=0.02$). There were no significant differences in years of education ($t(26.436)=-1.31$ $p=0.20$) or MoCA scores ($t(29.676)=0.089$, $p=0.93$) between males and females at baseline. Males had a significantly longer follow-up time compared to females ($t(34.736)=-2.039$, $p=0.049$). From baseline to follow-up, MoCA scores significantly declined by an average of 1.05 points ($t(38)=3.241$, $p=0.002$).

The results of the TBSS analysis comparing DTI metrics from baseline to follow-up for all study participants revealed no statistically significant changes in FA. However, there were statistically significant changes in MD. Decreases in MD were found in the right superior and inferior longitudinal fasciculus (including temporal part), right corticospinal tract, and the right inferior fronto-occipital fasciculus (see *Figure 3.1*). In the analysis comparing DTI metrics for males and females separately, stability in FA and MD was observed for both sexes. Individual changes in FA and MD values are displayed in spaghetti plots in *Figure 3.2* and *Figure 3.3,* respectively.

Table 3.1 *Descriptive Statistics of Participants Selected from the PPMI Healthy Control Cohort*

		All Participants (n=39)		Sex			
				Female (n=16)		Male (n=23)	
Variable	Timepoint	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (years)[a]	Baseline	63.66	7.33	60.51	5.96	65.50	7.53
	Follow-up	64.73	7.33	61.55	5.93	66.57	7.52
MoCA Total Score[b]	Baseline	28.23	1.09	28.25	1.18	28.21	1.04
	Follow-up	27.18	2.09	27.50	2.28	26.95	1.96
Whole brain FA	Baseline	0.462	0.018	0.459	0.015	0.463	0.021
	Follow-up	0.465	0.019	0.462	0.018	0.467	0.020
	Percent Change	0.688	3.069	0.615	2.844	0.740	3.279
Whole brain MD	Baseline	$7.1622e^{-4}$	$3.15e^{-5}$	$7.1635e^{-4}$	$2.79e^{-5}$	$7.1613e^{-4}$	$3.45e^{-5}$
	Follow-up	$7.0846e^{-4}$	$2.74e^{-5}$	$7.0818e^{-4}$	$1.84e^{-5}$	$7.0865e^{-4}$	$3.26e^{-5}$
	Percent Change	-1.003	3.348	-1.065	2.743	-0.961	3.772
Follow-up time (years)		1.07	0.14	1.02	0.08	1.10	0.16
Education (years)		15.72	2.74	15.00	3.14	16.21	2.37

Note. MoCA = Montreal Cognitive Assessment (maximum score = 30). FA = Fractional anisotropy. MD = mean diffusivity. [a]At baseline, males were significantly older than females ($t(38)=-2.47$, $p=0.02$). [b]MoCA scores significantly declined by an average of 1.05 points from baseline to follow-up ($t(38)=3.241$, $p=0.002$).

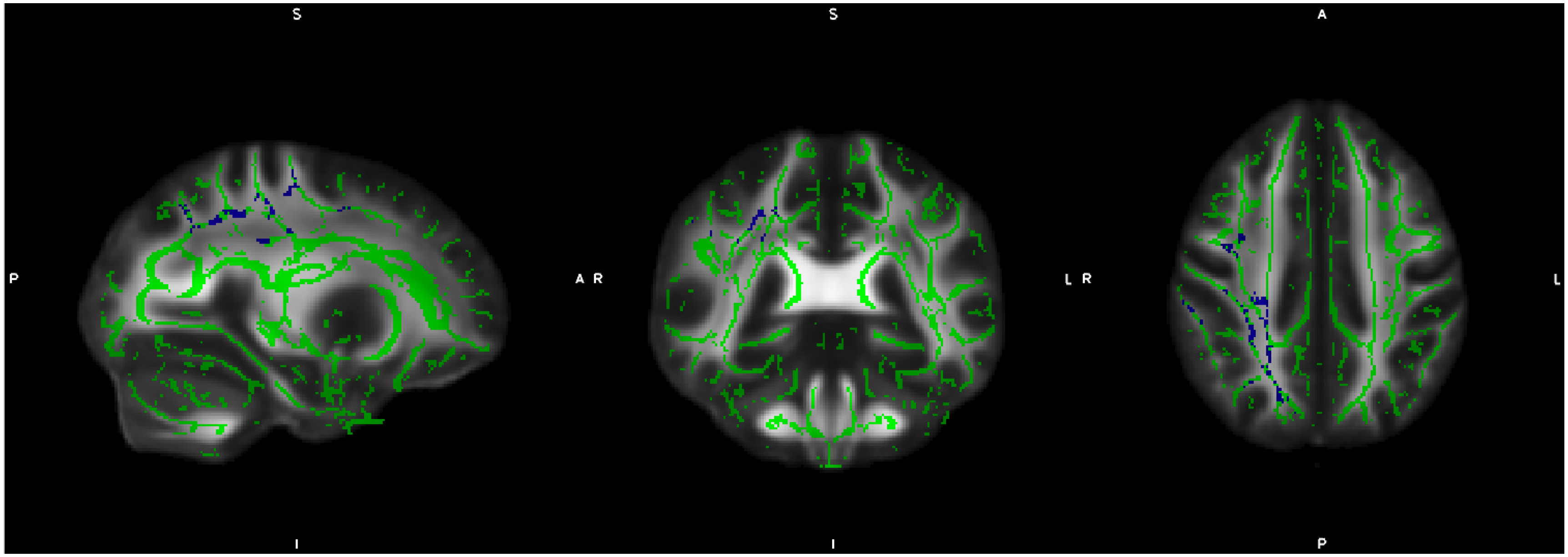

Figure 3.1 Decreases in MD (blue) from baseline to follow-up for all participants (*p*<.05).

These changes are seen in the right superior and inferior longitudinal fasciculus (including temporal part), right corticospinal tract, and the right inferior fronto-occipital fasciculus (JHU White Matter Tractography Atlas), which are overlaid on the mean white matter skeleton (green) on the standard FMRIB58_FA_1mm brain.

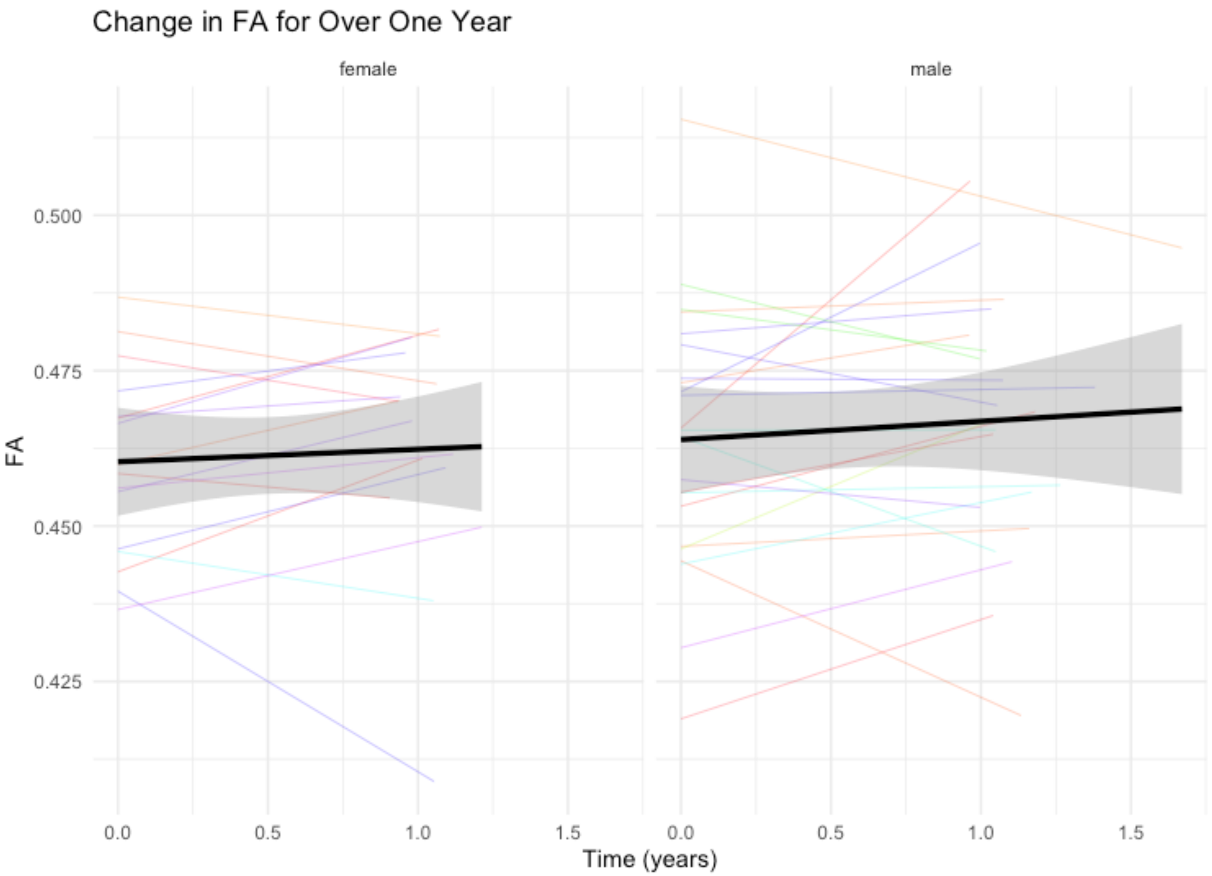

Figure 3.2 Changes in whole brain FA values for all participants from baseline to follow-up separated by sex.
FA values were calculated using the formula seen in Table 1.2 and fall between 0 and 1, where higher values represent greater anisotropy or directionality of diffusion, which correspond to greater white matter integrity. Each coloured line represents a single participant. The mean change is shown in the thick black line with the grey shaded representing the standard error of measurement. The TBSS analysis revealed no significant changes in FA from baseline to follow-up.

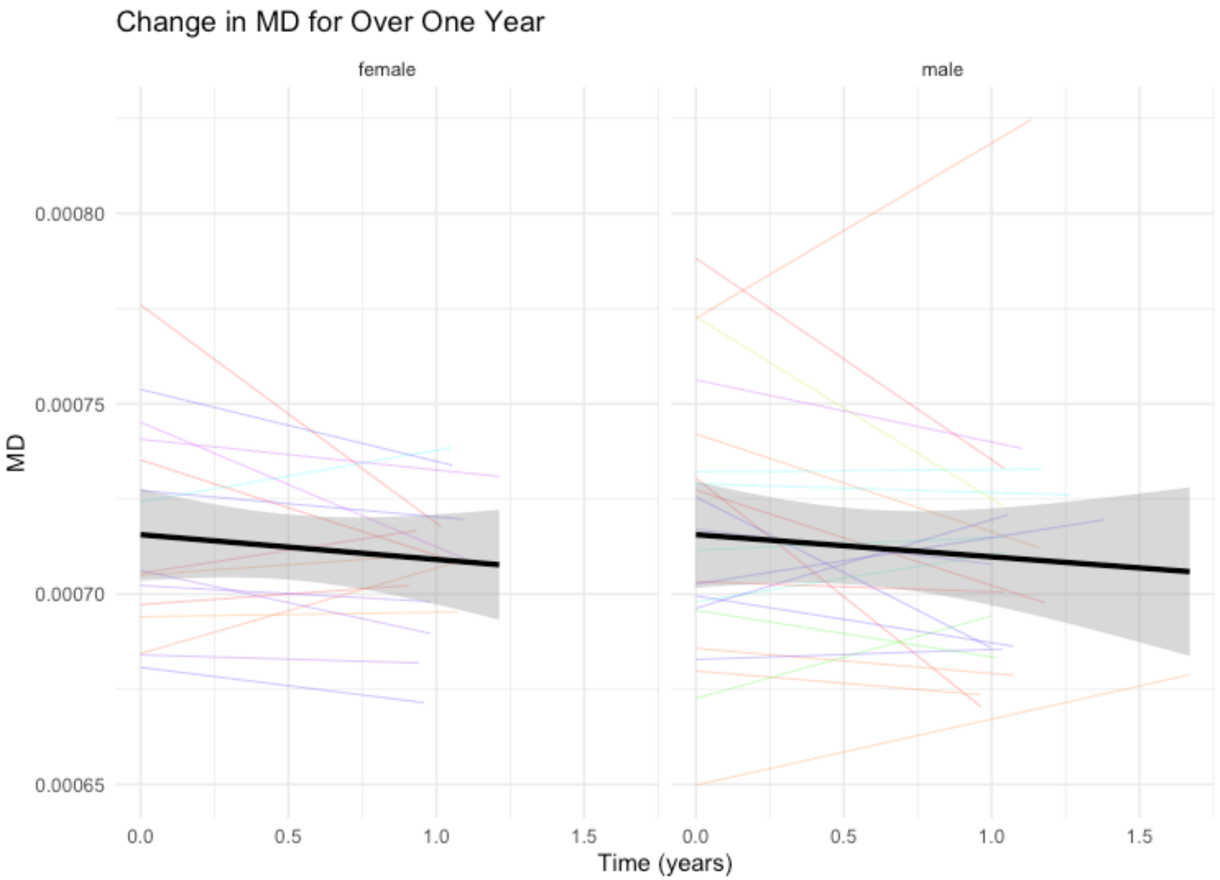

Figure 3.3 Changes in whole brain MD values for all participants from baseline to follow-up separated by sex.
MD was derived using the formula shown in Table 1.2. Each coloured line represents a single participant. The mean change is shown in the thick black line with the grey shaded representing the standard error of measurement. The TBSS analysis for all participants revealed decreases in MD from baseline to follow-up, while examining changes in MD separately by each sex revealed overall stability in whole brain MD values.

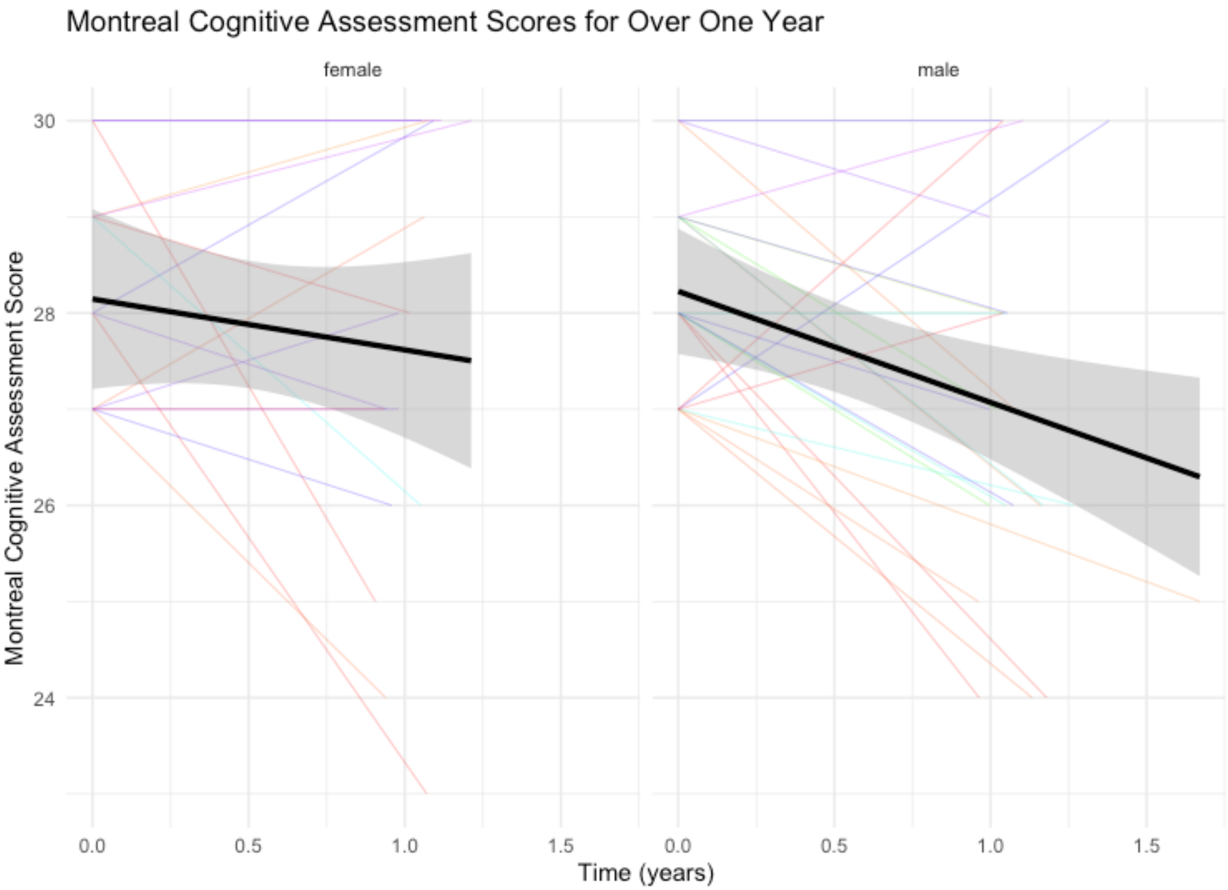

Figure 3.4 Changes in Montreal Cognitive Assessment Scores for females and males from baseline to follow-up.
Each coloured line represents a single participant. The mean change is shown in the thick black line with the grey shaded representing the standard error of measurement. MoCA scores significantly declined by an average of 1.05 points ($t(38)=3.241$, $p=0.002$) for all participants.

Discussion

In the current study, the expected longitudinal changes in white matter in the brain was investigated in a secondary TBSS analysis using DTI data for a sample of 39 healthy older adults selected from the healthy control group of the PPMI. For the primary research question of how white matter changes over time, it was hypothesized that stability or decline in white matter microstructural integrity would be observed over the approximately one-year study period (i.e., stability or decreased FA and stability or increased MD). This hypothesis was partially supported. Comparing the whole sample from baseline to approximately one year later, stability in whole brain FA was observed. However, contrary to what was hypothesized, decreased MD was found in the right superior and inferior longitudinal fasciculus, the right corticospinal tract, and the right inferior fronto-occipital fasciculus. These findings were unexpected, as previous longitudinal studies of brain changes among healthy older adults have found decreases in FA (Bender et al., 2016; Charlton et al., 2010; Lövdén et al., 2014a; Sexton et al., 2014) and increases in MD (Charlton et al., 2010; Lövdén et al., 2014b; Sexton et al., 2014). The finding of decreased MD generally reflects decreased diffusivity within a voxel, which suggests greater barriers to diffusion that are often interpreted as increased microstructural integrity of white matter. However, DTI metrics can be misleading, particularly in regions of the brain with multiple crossing white matter fibers of varying orientation (Figley et al., 2022). In this case, it can be helpful to examine other pieces of information. The small, but statistically significant decrease in MoCA scores across the entire sample were suggestive of very subtle decline (or even clinical stability) in overall cognitive ability among study participants. Thus, the possibility that the finding of decreased MD may have been a spurious finding cannot be ruled out.

Changes in white matter were further examined with a secondary research question as to whether white matter changes occur similarly or differently for healthy older adult males and females. It was hypothesized that males and females will show similar stability or decline over time, as previous research has been limited. The TBSS analysis was repeated separately for each sex, with the results demonstrating overall stability in FA and MD from baseline to follow-up for both males and females. This finding suggested that there are limited sex differences in white matter changes across one year among males and females, in agreement with a large review of the literature finding limited brain differences between males and females after controlling for differences in brain size (Eliot et al., 2021). Others have also found no differences in white matter microstructure and gender (not defined) (Ouyang et al., 2021). Nonetheless, it was unusual that the longitudinal decreases in MD that were found in the whole sample was not observed when examining either sex separately.

To further investigate the changes in white matter among the study sample, individual trajectories of change FA and MD over time were plotted. These plots demonstrated significant variability in FA and MD change over time for both males and females, such that some individuals demonstrated different changes in DTI metrics reflective of improvement, some reflected stability, and others showed decline in overall white matter microstructural integrity. These different trajectories may have created a "fanning" effect, such that the overall group trend of changes over time was ameliorated in the TBSS analyses. It will be important to for future research to identify factors that may predict these various trajectories among healthy older adults, with the goal to identify whether these trajectories can be used to predict future age-related decline in brain structure or function.

There are several limitations of the current study. First, the data used in this study was selected from the healthy control cohort of the PPMI. This limited the sample to 39 participants with DTI data available for both baseline and follow-up time points. There were a greater number of males than females, males were older than females by an average of approximately five years, and the time between the baseline and follow-up scan was approximately 0.08 years longer for males than for females, on average. PPMI did not collect gender-based data, so the results herein reflect similar stability in DTI metrics by biological sex assigned at birth, rather than one's social and self-ascribed gender identity. Further, the sample was largely Caucasian (one female participant identified as an ethnicity "Other" than White/Caucasian, Black, or Asian and one male participant identified as "Black"), which limits the generalizability of the findings. Future research should examine white matter changes in larger and more inclusive samples to investigate the intersection of biological and sociological determinants in healthy aging.

Second, the analysis method of TBSS is also not without limitations. Even when following the preprocessing pipeline appropriately, there can be anatomic inaccuracies and bias in the skeleton projection step, which can potentially cause problems when analyzing small tubular structures due to inter-subject differences in its location and crossing or kissing fibres that can alter DTI metrics (Bach et al., 2014). Further, the interpretation of DTI metrics is complex. Changes in DTI metrics can be attributed to partial volume effects due to characteristics in fiber bundles such as thickness, orientation, and curvature (Vos et al., 2011). Additionally, raw DTI metric values were plotted for the purposes of exploring general trends in these data. It is important to recognize that a single value was used to quantify the average FA or MD across all white matter tracts in the brain for each participant at each time point. During this process, the specific spatial information regarding white matter microstructural integrity within brain regions was lost. Future research could continue to investigate longitudinal changes in white matter by extracting raw values within specific regions of interest.

Nonetheless, these findings contribute to a growing body of literature focused on documenting the expected changes in white matter that occur as part of the normal healthy aging process among older adults. Overall stability in FA and decreases in MD were found in this sample of 39 participants healthy control participants from the PPMI. Stability in FA was in line with the hypotheses, but the finding of decreased MD was unexpected. Due to the limited sample size, known limitations of TBSS, and the significant variability among individual trajectories of white matter change over time, it is possible that the finding of decreased MD may have been a spurious finding. As such, it will be important to continue to investigate white matter changes among healthy older adults in larger samples. Conducting secondary data analyses on healthy control groups from large neuroimaging studies such as the PPMI is a convenient and cost-effect way to replicate these findings. It will also be important to include sex and gender-based analyses in future studies of healthy aging, which could be relevant information for the etiology of age-related diseases that differ by sex and gender.

Chapter 4: Study 3 - Sex Similarities and Differences in Cognition: A Longitudinal Study of Healthy Control Participants from the Parkinson Progression Markers Initiative

Abstract

Background: As the global population of older adults continues to rise, there is an urgent need to understand the expected cognitive changes that occur during healthy aging. Despite sex-based differences in age-related diseases and life expectancy, limited research has explicitly examined sex differences in aging. Longitudinal study designs are particularly important because they allow for examination of intraindividual change over time as a function of sex.

Study Aims: The objective of the current study was to answer several primary questions:

1) What are the cognitive changes that occur over time in a sample of healthy adults?
2) Are there sex differences in cognition among healthy adults?
3) Are there sex differences in rates of change in cognition over time?

Method: The current study utilized longitudinal data from the healthy control group of the Parkinson's Progression Markers Initiative to examine cognitive changes in healthy adult males (n=125, mean age=61.61, SD=10.97) and females (n=68, mean age=59.44, SD=11.56). Participants completed neuropsychological measures annually for up to five years (i.e., six time points). These measures included the Montreal Cognitive Assessment, Letter Number Sequencing, Semantic Fluency, Symbol Digit Modalities Test, Benton Judgment of Line Orientation, Hopkins Verbal Learning Test-Revised Immediate and Delayed Recall. Neuropsychological measures were modeled separately as a single outcome variable in a series of 2-level linear mixed models. Within-person changes in cognition over time (i.e., level-1) and between-group differences longitudinal change trajectories as predicted by sex (i.e., level 2) were examined in a hierarchical fashion, with the best fitting models including random intercepts and random slopes. Effects of age and education were also examined.

Results: Overall, there was no clinically meaningful systematic change in neuropsychological scores over time. Significant sex differences on neuropsychological scores were seen on the SFT, SDMT, BJLOT, and HVLT-R Immediate and Delayed Recall. However, rates of change in cognition over time did not significantly differ by sex. Higher baseline age predicted lower scores for all neuropsychological outcome measures, and higher education predicted higher scores for all neuropsychological outcome measures except for the MoCA. There were significant time by age interactions for the models predicting MoCA and BJLOT scores, although these estimates were exceedingly small and not clinically meaningful.

Conclusions: Stability in cognition was observed over a span of five years in a sample healthy control participants from a large multi-site longitudinal study. Although there were sex differences in certain domains of cognitive function, rates of cognitive change over time did not significantly differ by sex. Higher age and lower education were predictive of lower cognitive functioning. Great intraindividual variability in trajectories was observed within this study. Future research should continue to examine factors that predict individual trajectories of aging in healthy individuals.

Keywords (5): sex differences, aging, cognition, multi-level modelling, longitudinal

Introduction

Worldwide, the population of adults over the age of 65 is expected to double from 703 million to 1.5 billion by 2050 and the population of adults over the age of 80 is expected to triple from 143 million to 426 million by 2050 (United Nations et al., 2020). These shifts in population demographics emphasize the urgent need to understand normal age-related changes in cognition, to aid in promoting health and longevity into the later decades of life.

Most of the literature to date has examined cognitive aging without investigating differences in sex and gender. However, sex and gender are important factors that may influence cognitive aging trajectories, especially given sex-based differences in longevity and predisposition to neurodegenerative conditions (Zagni et al., 2016). One of the challenges in understanding sex differences is the use of inconsistent terminology and difficulties applying these concepts in research (Day et al., 2016). Several national and international organizations reached consensus that *sex* refers to biological and physical characteristics of humans whereas *gender* refers to socially constructed characteristics of personal identity. However, there is a bidirectional entanglement of these factors whereby it is extremely challenging to differentiate inborn biological sex factors from social and environmental gender learning (Eliot et al., 2021), such that some researchers argue that the composite term “sex/gender” be used (Kaiser, 2012; Rippon et al., 2014).

There has been great interest and controversy in the study of sex/gender differences in cognition since the 1970s (Maccoby & Jacklin, 1974). Although significant sex/gender differences have been found in specific cognitive abilities, including visuospatial working memory (Voyer et al., 2017), mental rotation (Maeda & Yoon, 2013), episodic memory

(Asperholm et al., 2019), and verbal abilities (Barel & Tzischinsky, 2018; Zhang et al., 2017) the effect sizes of these differences are usually quite small. The "gender similarities hypothesis" posits that males and females are largely similar on most, but not all variables (Hyde, 2005). Although males and females are more similar cognitively than they are different, questions remain about sex/gender differences in the normal aging trajectory. For example, do healthy older adult males and females experience similar changes in cognitive abilities over time? If there are differences in rates of change, in which domains do they occur, and for which sex?

Overall, most longitudinal studies have found minimal sex/gender differences in the rate of cognitive changes over time in older healthy adults. A meta-analysis of 13 longitudinal studies concluded that sex/gender was not a determining factor in cognitive decline amongst individuals aged 60-80 years (Ferreira et al., 2014). Although the meta-analysis did not find sex/gender differences in overall cognitive level, females outperformed males on tests of episodic memory (de Frias et al., 2006; Wetherell et al., 2002) and males outperformed females on tests of visuospatial ability (de Frias et al., 2006; Lovden et al., 2004; Wetherell et al., 2002). Sex/gender differences for tests of semantic memory and attention were more inconsistent (Ferreira et al., 2014). The overall evidence suggested that there are few if any sex/gender differences in cognitive aging in healthy older adults before age 80 and that other factors, such age, education, and social behaviour were found to be more predictive of cognitive changes than sex/gender (Ferreira et al., 2014). However, other studies, such as the Baltimore Longitudinal Study on Aging, have found steeper rates of decline for males on measures of mental status, perceptuomotor speed and integration, and visuospatial ability, suggesting that females may have a higher resilience to age-related cognitive decline compared to males (McCarrey et al., 2016). Specifically, females showed higher baseline scores on tests of verbal learning and memory, fluent language production, mental status, and psychomotor speed, while males showed an advantage on tests of visuospatial ability. Both males and females showed a decline on all cognitive measures over the study period of 3-9 years (McCarrey et al., 2016).

More research is needed to examine trajectories of aging as a function of sex and longitudinal study designs are necessary for measuring intraindividual change over time. The Parkinson's Progression Markers Initiative (PPMI) is an ongoing observational longitudinal study designed to investigate biomarkers related to Parkinson's disease. For up-to-date information, please see: https://www.ppmi-info.org/. The PPMI study presents an opportunity to investigate the expected aging trajectory among healthy older adults by using data from the healthy control cohort. Biological sex difference can also be investigated using this data; however, no information based on gender was collected. The objective of the present study was to use these data to investigate 1) age-related changes in cognition among healthy older adults, 2) sex differences in cognition, and 3) sex differences in rates of cognitive change over time. Multi-level mixed modelling poses several advantages for longitudinal data analysis, including flexibly handling missing data and capturing the intraindividual variance between measurement occasions. Using performance on neuropsychological measures as the dependent variables administered annually for up to six time points over five years, a series of 2-level linear mixed models were used to answer three primary research questions and hypotheses:

1) What are the cognitive changes that occur among healthy adults over a five-year time frame (i.e., level-1, within-person change)?

It was hypothesized that most individuals will be stable or show small decreases in scores on the neuropsychological measures over time. Based on previous literature, it was expected that declines in processing speed, verbal fluency, and some aspects of executive functioning and memory may be observed (Ferreira et al., 2014), while vocabulary may remain more stable (Harada et al., 2013).

2) Are there sex differences in cognition (i.e., level-2, differences in intercepts)?

It was hypothesized that there would likely be some sex differences in cognition. Consistent with previous research (de Frias et al., 2006; Wetherell et al., 2002), females were predicted to score higher on verbal measures (i.e., Semantic Fluency Test, Hopkins Verbal Learning Test - Revised) and males were predicted so score higher on visuospatial measures (i.e., Judgment of Line Orientation.

3) Are there sex differences in rates of change in cognition over five years (i.e., level-2, between-group differences in change, differences in slopes)?

Given the subtle sex differences in verbal fluency that have been found in previous research, it was hypothesized that men may experience greater decline on the verbal fluency relative to women. Other domains were predicted to change similarly between the sexes.

The current study contributes to a growing body of literature aimed at investigating expected cognitive changes due to normal aging and how these changes may be similar or different by biological sex.

Methods

Participants

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org. PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, which can be found at www.ppmi-info.org/fundingpartners**.** Participants from the PPMI were eligible for inclusion in the present study if: a) they were classified as a healthy control group participant by PPMI, and b) neuropsychological data was available for at least one time point. Initial inclusion and exclusion criteria used for the healthy control cohort by the PPMI investigators at the baseline time point is found in Appendix A. Up to six annual time points were available (i.e., baseline or time 0, time 1, 2, 3, 4, 5). Participants were excluded from the present analysis if PPMI's criteria for a healthy control participant was not met at any time point, which PPMI defined as two or more neuropsychological test scores < 1.5 standard deviations below the mean or PPMI investigator rating of mild cognitive impairment or dementia.

It is important to note that the PPMI Operations Manual indicated that a sex/gender variable was based on "gender at birth". Based on the consensus definitions in the literature, it is apparent that PPMI's use of "gender" corresponds with the consensus definitions of "sex" and this variable was treated as such. Unfortunately, PPMI did not collect data corresponding to the consensus definitions of "gender", so it was not possible to examine the effects of gender within this database. Ethical approval was for original data collection was obtained by the PPMI investigators, and ethical approval for secondary data analysis for the present study was also obtained from the University of Victoria Ethics Board. These criteria yielded a total of 193 participants (68 female) who were measured across one to six occasions. Participants had an average of 3.34 measurement occasions.

Measures

Neuropsychological data were collected by PPMI researchers at baseline (time 0) and then approximately annually for five years, for up to a total of six annual time points (i.e., baseline or time 0, 1, 2, 3, 4, 5). The present study utilized scores from all available neuropsychological measures administered to the healthy control cohort. This yielded six neuropsychological measures included the Montreal Cognitive Assessment (MoCA), the Letter Number Sequencing test (LNS), the Semantic Fluency Test (SFT), the Symbol Digits Modalities Test (SDMT), the Benton Judgement of Line Orientation Test (BJLOT), and the Hopkins Verbal Learning Test – Revised (HVLT-R). A brief description of each measure and the specific scores used as outcome measures is summarized in Table 1.

Montreal Cognitive Assessment (MoCA).

The MoCA is a brief screening measure for cognitive dysfunction that takes approximately 10 minutes to administer (Nasreddine et al., 2005). This measure is often used a screening measure of global cognition, as it includes brief tasks assessing attention and concentration, executive function, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. The total score is out of 30 points. Scores below 25 are suggestive of cognitive impairment. Test-retest reliability ranges from 0.42-0.48 and reliable change indices for healthy older adults indicate a change of more than 4 points to be clinically significant (Kopecek et al., 2017).

Letter Number Sequencing (LNS).

The LNS test is a measure of working memory from the Wechsler Adult Intelligence Test - Third Edition. Participants alphabetically and numerically sequenced a random string of letters and numbers. One point was awarded for each correctly ordered string of letters and numbers. Seven items were administered, which increased from a string of two for the first item (i.e., one letter and one number) to a string of eight (i.e., four letters and four numbers) for the seventh item. Each item had three trials of the same number of letters and numbers. The test was discontinued after a score of 0 for three trials of the same item, for a possible total score of 21.

Semantic Fluency Test (SFT).

The SFT test is a measure of verbal fluency and memory retrieval (Goodglass & Kaplan, 1972). Participants were asked to name as many items as possible from a certain category. The test included three trials (i.e., animals, vegetables, fruits) of one minute each. Scores reflect the total number of words from all three trials.

Symbol Digit Modalities Test (SDMT).

The SDMT is a measure of processing speed, complex attention, working memory, and visual scanning (Smith, 1973). Participants have 90 seconds to write numbers that correspond with symbols by referring to a number to symbol key. Form 1 was used for the baseline assessment and then alternated with Form 2 for each measurement occasion. Scores reflect the total correct items completed.

Benton Judgement of Line Orientation (BJLOT).

The BJLOT (Benton, 1994) is a measure of visuospatial judgment and orientation. Participants were to match lines presented on different angles to a standard array of several lines arranged from 0 to 180 degrees. The standard version of the test is scored out of 30. However, half of the test items were administered at each measurement occasion by alternating even and odd items (i.e., the total raw score is out of 15) as a way of limiting practice effects.

Hopkins Verbal Learning Test – Revised (HVLT-R).

The HVLT-R (Brandt & Benedict, 2001) is a measure of verbal learning and memory. Participants were read a list of 12 words (four words from three different semantic categories) and asked to immediately recall the word list over three learning trials. The participants were then asked to freely recall the word list after a 20–25-minute delay. A recognition trial was also administered. Variables used in the present study were the sum of the number of words recalled over the three learning trials (i.e., HVLT-R immediate recall) and the total words recalled after the 20-25-minute delay (i.e., HVLT-R delayed recall). Six different forms (i.e., unique word lists) were available, and a unique one was used for each measurement occasion.

Table 4.1 *Description of Neuropsychological Measures and Cognitive Domains Assessed*

Neuropsychological Measure	Cognitive Domain(s) Assessed	Description of Measure	Score Calculation	Repeated Measurements
Montreal Cognitive Assessment (MoCA)	Overall cognitive ability	A screener for mild cognitive impairment and dementia. Brief assessment of visuospatial/executive, naming, memory, attention, language, abstraction, and orientation.	Sum of correct items (maximum = 30)	Same form each time point.
Letter Number Sequencing (LNS)	Working memory	Sequence a string of letters and numbers of increasing length alphabetically and numerically.	Sum of correctly sequenced trials (maximum = 21)	Same form each time point.
Semantic Fluency Test (SFT)	Verbal fluency, executive functioning, and memory retrieval	Name as many items in a category in one minute.	Sum of three categories (no maximum)	Same categories each time point.
Symbol Digit Modalities (SDMT)	Executive function, working memory, and processing speed	Write numbers to match symbols using a key.	Number of correctly coded symbols. Maximum = 110	Alternated between Form 1 and Form 2 for each time point.
Benton Judgement of Line Orientation (JOLO)	Visuospatial ability	Judge the angle of lines by matching them to a key.	Number of correctly identified line orientations (maximum = 15)	Alternated between odd and even items at each time point.
Hopkins Verbal Learning Test (HVLT)	Verbal learning and memory	Learn a list of unrelated words over three learning trials, then recall it freely after a delay.	Immediate Recall is the sum of three learning trials (maximum score = 36) Delayed Recall is the sum of correctly recalled words after a 20-minute delay (maximum score = 12)	Unique form for each time point (i.e., 6 unique word lists)

Study Design

A series of 2-level linear mixed models were created using each of the seven neuropsychological scores as a single dependent or outcome variable in seven separate models (i.e., MoCA, SDMT, LNS, SFT, BJLOT, HVLT-R immediate and delayed). These models were used to estimate within-person change in cognition over time (i.e., level-1, repeated measures or time level), between-person differences in cognition by sex, and between-person differences change trajectories as predicted by sex (i.e., level-2, individual level). The structure of the 2-level design is displayed in Figure 1, where repeated measures are nested within individuals. Additional level-2 time-invariant predictors were also investigated, including baseline age and education. Model specification and equations are shown in Table 2. A detailed description of the multi-level model specification and equations follows.

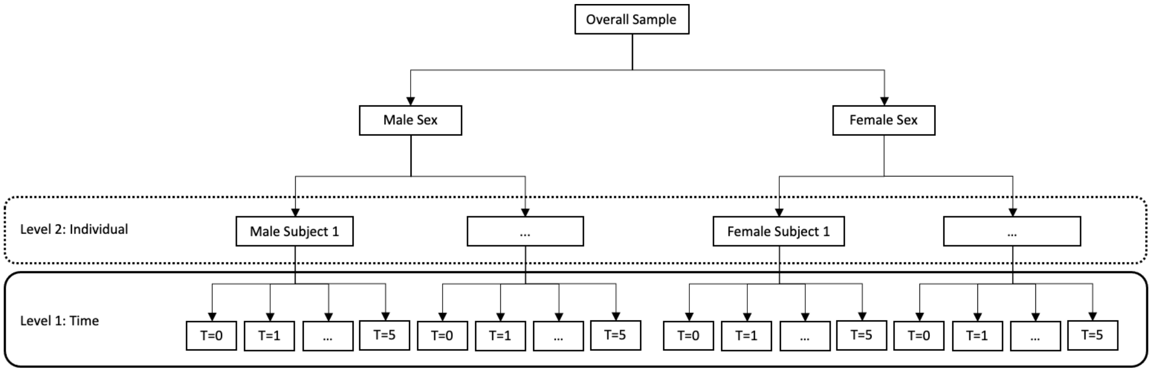

Figure 4.1 Schematic of the 2-level linear mixed model.

Each neuropsychological score was treated as a dependent variable and modelled separately. Level-1 represented within-person repeated measurements in cognition for up to six time points and level-2 represented between-person differences based on several time-invariant predictors. This figure displays the primary focus of the present study on evaluating between-person differences in cognitive change over time as predicted by the categorical predictor of sex (i.e., male or female). Other time-invariant level-2 predictors investigated were years of education and age at baseline.

Model Specification and Equations

Unconditional Model: Random Intercept Only.

First, an unconditional model with a random intercept only was created. This model is often referred to as the "unconditional model" or "null model" as it is the most basic model possible without the inclusion of any predictor variables. Subsequent models with predictor variables were compared to this unconditional model to determine whether additional predictor variables explained additional variance in the data. The model is composed of two levels. The level 1 (repeated measures or time level) model is as follows:

$$NP_{ti} = \beta_{0i} + e_{ti} \tag{1}$$

This equation shows the neuropsychological score for the i-th participant at measurement occasion t. Where NP denotes the dependent variable (i.e., score on the neuropsychological measure), t represents the six different measurement occasions (i.e., baseline, time 1, time 2, time 3, time 4, and time 5 follow up points), and i represents the individual being measured (i.e., participant 1 up to total number of participants for that specific neuropsychological measure). e_{ti} is the within-person random error which captures the difference between the observed NP score at time t for the i-th participant and the predicted (average) score of the i-th participant. e_{ti} is assumed to be normally distributed with variance equal σ^2 (i.e., $e_{ti} \sim N(0, \sigma^2)$). This parameter captures the within-person variation. β_{0i} is the estimated NP score for the i-th individual, which is further defined in level-2 (individual level) of the model:

$$\beta_{0i} = \gamma_{00} + U_{0i} \tag{2}$$

γ_{00} represents the grand mean of the participants' average NP scores (i.e., mean of the individual participant means) and U_{0i} is the difference between the i-th individual's mean NP score and the grand mean. The assumption is made that U_{0i} is normally distributed with a variance equal to τ_{00} (i.e., $U_{0i} \sim N(0,\ \tau_{00})$). The within-person random errors (e_{ti}) and the between-person random effects (U_{0i}) are assumed to be independent.

The random intercept model was created for each dependent variable (i.e., each neuropsychological measure). This is shown below by substituting the level-2 individual model (Equation 2) into the level-1 repeated-measures model (Equation 1) to create the composite random intercept only model (Equation 3):

$$NP_{ti} = [\boldsymbol{\beta_{0i}}] + e_{ti} \tag{1}$$

$$NP_{ti} = [\boldsymbol{\gamma_{00}} + \boldsymbol{U_{0i}}] + e_{ti} \tag{3}$$

As discussed, this model includes no predictors. It tests whether the intercept, γ_{00}, which is a fixed effect (i.e., the same for all individuals) and reflects the mean of the individual means of the dependent variable, is significantly different than 0. This is useful for calculating the intra-class correction (ICC), which represents the magnitude of dependency between observations using the within-person variances (σ^2) and between-person variances (τ_{00}). The ICC is also sometimes referred to as the Variance Partition Coefficient (VPC), as it provides a value between 0 and 1, and essentially represents the proportion of variance due to between-person variance relative to the total variation due to the between and within-person variance (Equation 4).

$$Intraclass\ Correlation\ (ICC)\ or\ Variance\ Partition\ Coefficient = \frac{\tau_{00}}{\tau_{00} + \sigma^2} \tag{4}$$

Values closer to 0 indicate a greater proportion of within-person variance and values closer to 1 would indicate a greater portion of between-person variance. In other words, the ICC or VPC represents the degree of dependency or correlation of scores on the outcome variable within the same individual across their individual measurement occasions. Higher values indicate more highly correlated scores within individuals over time, representing less variation within-persons over time and greater between-person variance over time.

Unconditioned Growth Model: Fixed Linear Time and Random Intercept.

Next, a single predictor variable, time, was added into the random intercept only model. These linear growth models were created for each of the neuropsychological outcome measures using the following Level-1 (repeated-measures level) model:

$$NP_{ti} = \beta_{0i} + \beta_{1i}TIME_{ti} + e_{ti} \tag{5}$$

This linear growth model introduced a random slope into the equation, β_{1i}, that varied for every i-th individual. Notably, although the term "growth" implies increasing values, here it is used as a statistical term to represent how values change over time (i.e., growth, decline, or stability). The repeated measures aspect of the model is represented by $TIME_{ti}$, which was coded so that the baseline assessment time was equal to 0 for all participants. Subsequent follow-up assessments for each participant were unique and calculated in exact years since the baseline assessment. Coding time this way facilitated interpretation of the model, such that β_{0i} (i.e., the intercept) was equal to the estimated NP score for the i-th individual at the baseline assessment (i.e., $TIME_{ti}$ = 0) and β_{1i} (i.e., the slope) was equal to the estimated average yearly change in NP score for the i-th individual. As described above in the discussion of the intercept only model, e_{ti} is the within-person random error and was assumed to be normally distributed with variance equal σ^2 (i.e., $e_{ti} \sim N(0, \sigma^2)$), which captured the within-person variation. The regression coefficients from each individual growth model (i.e., the intercept, β_{0i}, which represents the NP score at the baseline assessment, and the slope, β_{1i}, which represents the average yearly change in NP score), were summarized by two equations, the level-2 (individual level) models:

$$\beta_{0i} = \gamma_{00} + U_{0i} \tag{6}$$

$$\beta_{1i} = \gamma_{10} + U_{1i} \tag{7}$$

γ_{00} is the average NP score at the baseline assessment (i.e., $TIME_{ti} = 0$) across the total number of participants and γ_{10} is the average yearly change in NP score across the total number of participants. U_{0i} and U_{1i} are the between-person random effects and were both assumed to be normally distributed. U_{0i} is the difference between the individual intercept (i.e., β_{0i}, the estimated NP score for the i-th individual at the baseline assessment where $TIME_{ti}$ = 0) and the average intercept (i.e., γ_{00}, the average NP score at the baseline assessment where $TIME_{ti} = 0$). U_{1i} is the average difference between the estimated yearly change in NP score for the i-th individual (i.e., β_{1i}) and the average yearly change in NP score across all the participants (i.e., γ_{10}). A composite model was created by substituting the values for the individual intercepts (i.e., β_{0i}, as shown in Equation 6) and individual slopes (i.e., β_{1i}, as shown in Equation 7) into the level-1 repeated measures linear growth model (i.e., introduction of both a random intercepts and random slopes, as shown in Equation 5). The substitutions are displayed in bold and in brackets, as follows:

$$NP_{ti} = [\boldsymbol{\beta_{0i}}] + [\boldsymbol{\beta_{1i}}]TIME_{ti} + e_{ti} \tag{5}$$

$$NP_{ti} = [\boldsymbol{\gamma_{00}} + \boldsymbol{U_{0i}}] + [\boldsymbol{\gamma_{10}} + \boldsymbol{U_{1i}}]TIME_{ti} + e_{ti} \tag{8}$$

Expanding the equation provides the following result:

$$NP_{ti} = \boldsymbol{\gamma_{00}} + \boldsymbol{U_{0i}} + \boldsymbol{\gamma_{10}}TIME_{ti} + \boldsymbol{U_{1i}}TIME_{ti} + e_{ti} \tag{9}$$

Conditional Growth Model.

Next, a time-invariant predictor, sex, was added to the model. Adding this predictor afforded the same equation for the level-1 repeated measures model with time as level-1 predictor:

$$NP_{ti} = [\boldsymbol{\beta_{0i}}] + [\boldsymbol{\beta_{1i}}]TIME_{ti} + e_{ti} \tag{5}$$

The time-invariant predictor of sex was added at the level-2 (individual level) models:

$$\beta_{0i} = \gamma_{00} + \gamma_{01}SEX_1 + U_{0i} \tag{10}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}SEX_1 + U_{1i} \tag{11}$$

SEX_1 was categorical and coded as 0 for females and 1 for males. Substituting the level-2 individual models (Equations 10 and 11) with a time-invariant predictor of SEX_1 into the level-1 repeated measures model (Equation 5) resulted in the following equation (with the substitution shown in bold and in brackets):

$$NP_{ti} = [\boldsymbol{\beta_{0i}}] + [\boldsymbol{\beta_{1i}}]TIME_{ti} + e_{ti} \tag{5}$$

$$NP_{ti} = [\boldsymbol{\gamma_{00} + \gamma_{01}SEX_1 + U_{0i}}] + [\boldsymbol{\gamma_{10} + \gamma_{11}SEX_1 + U_{1i}}]TIME_{ti} + e_{ti} \tag{12}$$

Expanding the equation afforded the following composite model:

$$NP_{ti} = \boldsymbol{\gamma_{00}} + \boldsymbol{\gamma_{01}SEX_1} + \boldsymbol{U_{0i}} + \boldsymbol{\gamma_{10}} * TIME_{ti} + \boldsymbol{\gamma_{11}SEX_1} * TIME_{ti} + \boldsymbol{U_{1i}} * TIME_{ti} + e_{ti} \tag{13}$$

Which can be rearranged to show the fixed effects (bolded) followed by the random effects (not bolded):

$$NP_{ti} = \boldsymbol{\gamma_{00}} + \boldsymbol{\gamma_{01}SEX_1} + \boldsymbol{\gamma_{10} * TIME_{ti}} + \boldsymbol{\gamma_{11}SEX_1 * TIME_{ti}} + U_{0i} + U_{1i} * TIME_{ti} + e_{ti} \tag{14}$$

Where γ_{00}, $\gamma_{01}SEX_1$, $\gamma_{10} * TIME_{ti}$, and $\gamma_{11}SEX_1 * TIME_{ti}$ are the fixed effects and represent the average model. The random effects are U_{0i}, $U_{1i} * TIME_{ti}$, and e_{ti}, where U_{0i} and $U_{1i} * TIME_{ti}$ represent the variation between the individual regression models and the average model and e_{ti} represents the variation between individual observations and the regression model for each individual participant.

Additional level-2 time-invariant predictors of baseline age and education were added to the conditional model above. Both predictors were "centered" by subtracting the group mean from each participant's baseline value so that the intercept of the model is interpreted as the estimated value in the outcome value for those with values of the predictor (i.e., age or education) equal to the group average. Finally, the following interaction terms were also incorporated into the final model: a) sex by time, and b) baseline age by time. A summary of the multi-level model specifications and equations are displayed in Table 2. The final model included random intercepts and random slopes.

Table 4.2 *Multi-level Model Specification Equations to Predict Neuropsychological Outcome Measures*

	Model		
Components	Unconditional: Random intercept only model	Unconditional Growth	Conditional Growth
Level-1 Equation	$NP_{ti} = \beta_{0i} + e_{ti}$	$NP_{ti} = \beta_{0i} + \beta_{1i}TIME_{ti} + e_{ti}$	$NP_{ti} = \beta_{0i} + \beta_{1i}TIME_{ti} + e_{ti}$
Level-2 Equation (s)	$\beta_{0i} = \gamma_{00} + U_{0i}$	$\beta_{0i} = \gamma_{00} + U_{0i}$ $\beta_{1i} = \gamma_{10} + U_{1i}$	$\beta_{0i} = \gamma_{00} + \gamma_{01}X_1 + U_{0i}$ $\beta_{1i} = \gamma_{10} + \gamma_{11}X_1 + U_{1i}$
Composite Equation	$NP_{ti} = \gamma_{00} + U_{0i} + e_{ti}$	$NP_{ti} = \gamma_{00} + U_{0i} + \gamma_{10}TIME_{ti} + U_{1i}TIME_{ti} + e_{ti}$	$NP_{ti} = \gamma_{00} + \gamma_{01}X_1 + \gamma_{10} * TIME_{ti} + \gamma_{11}X_1 * TIME_{ti} + U_{0i} + U_{1i} * TIME_{ti} + e_{ti}$
R-syntax: lmer(y~…	1 + (1 | ID)	time + (time | ID)	time * X_1+ (time | ID)
Predictors: Level-1		$TIME_{ti}$ = time in years between measurement occasions (continuous; centered on the baseline assessment), where t=exact time in years since the baseline assessment for individual participant, i. Data were unbalanced, meaning participants had 1-6 measurement occasions.	
Predictors: Level-2			X_1= sex (categorical; 0=female, 1=male) X_2= baseline age (continuous; centered on mean baseline age) X_3= years of education (continuous; centered on mean education)
Interactions			Time*Sex Time*Age

Note. Dependent variables were scores on neuropsychological measures (NP), which were modeled separately: Montreal Cognitive Assessment - Total Score /30; Letter Number Sequencing - Total Score /21; Semantic Fluency - Total Score (Sum of Animals, Vegetables, and Fruits); Symbol Digit Modalities Test - Total Score; Hopkins Verbal Learning Test – Revised Immediate Recall (Sum of Learning Trials 1-3) /36 & Delayed Recall /12; Benton Judgement of Line Orientation Test /15.

Procedure and Data Analysis

Data for the healthy control cohort was retrieved from the PPMI website, specifically from the file labeled "PPMI_Original_Cohort_BL_to_Year_5_Dataset_Apr2020.csv". Data analysis was conducted using the statistical software R and R Studio. The R package for Linear Mixed Effects Models using 'Eigen' and S4 (i.e., "lme4") was utilized for the multi-level linear modeling (Bates et al., 2022). The models were built using the R code shown in Appendix B. Models were then compared using the anova command from the "lmerTest" package for R, which produced type I, II, and III anova tables for fixed-effect terms with Satterthwaite and Kenward-Roger methods for denominator degrees of freedom for F-tests. Akaine Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were examined to determine the final model fit for the data.

Results

Descriptive statistics including the number of participants, sex, and race/ethnicity at each timepoint, along with the means and standard deviations for the participants' ages, years of education, neuropsychological test scores are shown in Table 3. Spaghetti plots were graphed to show the neuropsychological scores for each participant across all timepoints using the R package "ggplot2" (see Figures 2-8). Multilevel data analysis results are presented using the recommended Logical Explanations and Visualizations of Estimates in Linear mixed models (LEVEL) checklist (Monsalves et al., 2020). A summary of the final effect estimates for all seven neuropsychological outcome measures (which were modeled separately) is displayed in Table 4.4. Although the fit of the final models varied slightly depending on the neuropsychological outcome measure, the final models with all the predictors (i.e., sex, age, and education) and interactions of interest (i.e., time by sex, time by age) are shown. A detailed description of the specific model results for each neuropsychological outcome measure can be found in Appendix C.

Results for Question 1: What are the cognitive changes that occur among healthy adults over a five-year time frame (i.e., level-1, within-person change)?

All neuropsychological outcome measures, except for HVLT-R Immediate Recall, showed stability in scores within-individuals over time (i.e., no systematic change among all participants). On the HVLT-R Immediate Recall, individuals showed a slight improvement over time, by approximately 0.25 words per year.

Results for Questions 2: Are there sex differences in cognition (i.e., level-2, differences in intercepts)?

At baseline, females had higher scores on the SFT, SDMT, and the HVLT-R Immediate and Delayed Recall, while males had higher scores on the BJLOT. There were no significant sex differences on the MoCA or the LNS at baseline.

Baseline age was a significant predictor of cognition for all neuropsychological outcome measures, where higher age at baseline was reflective of lower cognitive performance.

Education was also a significant predictor in all models, except for the MoCA. Higher years of education was predictive of higher cognitive performance.

Results for Question 3: Are there sex differences in rates of change in cognition over five years (i.e., level-2, between-group differences in change, differences in slopes)?

Interaction of Time and Sex.

The interaction between change in neuropsychological scores over time and sex was non-significant in all models. Males and females did not systematically differ in their rates of change in any of the neuropsychological scores over time.

Interaction of Time and Age.

The interaction between change in neuropsychological scores over time and baseline age was non-significant for the LNS, SFT, SDMT, and the HVLT-R Immediate and Delayed. In the case of the MoCA and the BJLOT, the estimates were extraordinarily small, but indicated greater rates of decline in MoCA and BJLOT scores over time for each year above the baseline mean age of approximately 60 years. All other measures showed no differences in rates of neuropsychological scores over time based on age at the baseline assessment.

Table 4.3 *Participant Demographics and Descriptive Statistics*

		Time Point					
Mean (SD) or n		Baseline or 0	1	2	3	4	5
Years since baseline (SD)		NA	1.04 (0.08)	2.03 (0.10)	3.02 (0.09)	4.03 (0.09)	5.05 (0.18)
Total N		193	182	166	160	155	145
Race/Ethnicity	White	178	169	154	148	144	138
	Black	10	9	8	8	8	6
	Asian	1	1	1	1	0	0
	Other	4	3	3	3	3	1
Sex	Female	68	66	63	61	59	57
	Male	125	116	103	99	96	88
Age	Female	59.4 (11.6)	60.7 (11.5)	61.4 (12.0)	63.0 (11.8)	63.8 (12.0)	63.6 (11.3)
	Male	61.6 (10.9)	62.9 (10.9)	63.2 (11.2)	64.5 (10.7)	65.5 (10.6)	65.7 (10.5)
	Range	30.6-83.7	31.9-85.3	32.9-83.9	33.9-85.0	34.9-85.8	34.1-87.7
Education	Female	15.5 (2.7)	15.6 (2.7)	15.5 (2.7)	15.6 (2.8)	15.6 (2.8)	15.8 (2.8)
	Male	16.4 (2.9)	16.3 (2.9)	16.4 (2.9)	16.3 (3.0)	16.4 (3.1)	16.4 (3.0)
MoCA	Female	28.3 (1.1)	27.5 (2.2)	27.8 (3.9)	27.7 (2.1)	27.7 (2.4)	28.0 (2.1)
	Male	28.2 (1.1)	27.1 (2.2)	26.8 (2.5)	27.3 (2.1)	27.5 (2.3)	27.4 (2.2)
LNS	Female	10.9 (2.4)	10.9 (2.9)	11.1 (2.6)	11.1 (2.9)	11.3 (2.7)	11.5 (2.9)
	Male	10.8 (2.7)	11.0 (2.7)	10.9 (2.5)	11.0 (2.7)	10.9 (2.8)	10.9 (2.8)
SFT	Female	57.2 (10.7)	57.1 (9.8)	57.4 (11.5)	57.1 (11.0)	55.7 (10.5)	57.8 (11.4)
	Male	49.1 (10.5)	50.0 (11.2)	50.3 (11.3)	50.3 (11.5)	49.7 (11.8)	50.7 (11.7)
SDMT	Female	48.9 (10.2)	49.9 (10.1)	48.6 (10.2)	49.9 (9.8)	48.6 (10.0)	50.2 (10.3)
	Male	45.6 (10.7)	46.2 (11.4)	45.5 (11.2)	46.8 (11.8)	45.3 (11.7)	46.6 (12.3)
BJLOT	Female	12.4 (2.2)	11.6 (2.6)	12.4 (2.4)	11.7 (2.5)	12.4 (2.8)	11.9 (2.6)
	Male	13.5 (1.8)	13.2 (2.2)	13.5 (1.8)	13.1 (2.1)	13.2 (2.2)	13.3 (1.9)
HVLT-R							
Immediate Recall	Female	27.2 (4.4)	27.6 (4.5)	27.8 (4.8)	28.0 (4.8)	27.6 (4.5)	29.4 (3.8)
	Male	25.3 (4.4)	25.7 (4.7)	24.9 (4.8)	25.8 (5.3)	25.5 (5.4)	26.3 (5.8)
Delayed Recall	Female	9.9 (1.9)	9.7 (2.2)	10.1 (2.2)	9.6 (2.2)	9.9 (2.2)	10.4 (2.2)
	Male	8.9 (2.5)	8.8 (2.6)	8.7 (2.6)	9.0 (2.6)	8.7 (2.5)	9.4 (2.8)

Note. MoCA=Montreal Cognitive Assessment, LNS=Letter Number Sequencing, SFT=Semantic Fluency, SDMT=Symbol Digit Modalities Test, BJLOT=Benton Judgement of Line Orientation, HVLT= Hopkins Verbal Learning Test Revised, IR=Immediate Recall (Sum of 3 learning trials), DR=delayed recall.

Table 4.4 *Summary of the Multi-level Models to Predict Cognition on Seven Difference Neuropsychological Outcome Measures Over Five Years*

	Estimate (Confidence Interval)						
Outcome Measure	(Intercept)	Time	Sex	Age	Education	Time x Sex	Time x Age
MoCA	**27.92***** **(27.59 – 28.24)**	-0.04 (-0.15 – 0.06)	-0.24 (-0.65 – 0.16)	**-0.03***** **(-0.05 – -0.02)**	0.01 (-0.06 – 0.07)	-0.08 (-0.21 – 0.06)	**-0.01**** **(-0.01 – 0.00)**
LNS	**10.88***** **(10.36 – 11.39)**	0.05 (-0.05 – 0.16)	0.08 (-0.56 – 0.73)	**-0.08***** **(-0.11 – -0.05)**	**0.15**** **(0.05 – 0.25)**	-0.11 (-0.24 – 0.03)	0 (-0.01 – 0.00)
SFT	**57.04***** **(54.73 – 59.35)**	-0.04 (-0.43 – 0.36)	**-7.29***** **(-10.19 – -4.39)**	**-0.23***** **(-0.35 – -0.10)**	**0.51*** **(0.04 – 0.98)**	-0.02 (-0.53 – 0.48)	-0.01 (-0.03 – 0.01)
SDMT	**48.94***** **(47.01 – 50.87)**	0.06 (-0.31 – 0.42)	**-2.76*** **(-5.18 – -0.33)**	**-0.5***** **(-0.60 – -0.40)**	**0.71***** **(0.33 – 1.09)**	-0.05 (-0.51 – 0.42)	-0.01 (-0.03 – 0.01)
BJLOT	**12.2***** **(11.78 – 12.63)**	-0.03 (-0.12 – 0.06)	**1.21***** **(0.68 – 1.74)**	**-0.03*** **(-0.05 – -0.00)**	**0.15***** **(0.07 – 0.23)**	-0.03 (-0.14 – 0.09)	**-0.01*** **(-0.01 – -0.00)**
HVLT-R Immediate Recall	**27.17***** **(26.24 – 28.09)**	**0.25*** **(0.05 – 0.45)**	**-1.85**** **(-3.01 – -0.70)**	**-0.1***** **(-0.15 – -0.05)**	**0.28**** **(0.10 – 0.47)**	-0.19 (-0.44 – 0.07)	-0.01 (-0.02 – 0.00)
HVLT-R Delayed Recall	**9.77***** **(9.29 – 10.24)**	0.05 (-0.05 – 0.15)	**-0.93**** **(-1.52 – -0.33)**	**-0.05***** **(-0.08 – -0.03)**	**0.13**** **(0.04 – 0.22)**	-0.02 (-0.15 – 0.11)	0 (-0.01 – 0.00)

Note. MoCA=Montreal Cognitive Assessment, LNS=Letter Number Sequencing, SFT=Semantic Fluency, SDMT=Symbol Digit Modalities Test, BJLOT=Benton Judgement of Line Orientation, HVLT= Hopkins Verbal Learning Test Revised, IR=Immediate Recall (Sum of 3 learning trials), DR=delayed recall. Time, age, and education were centered so that time at baseline, mean baseline age of 60.62, and mean years of education of 16.07 were equal to 0. For sex, positive values indicate a male advantage and negative values indicate a female advantage. ***p<.001. **p<.01. *p<.05.

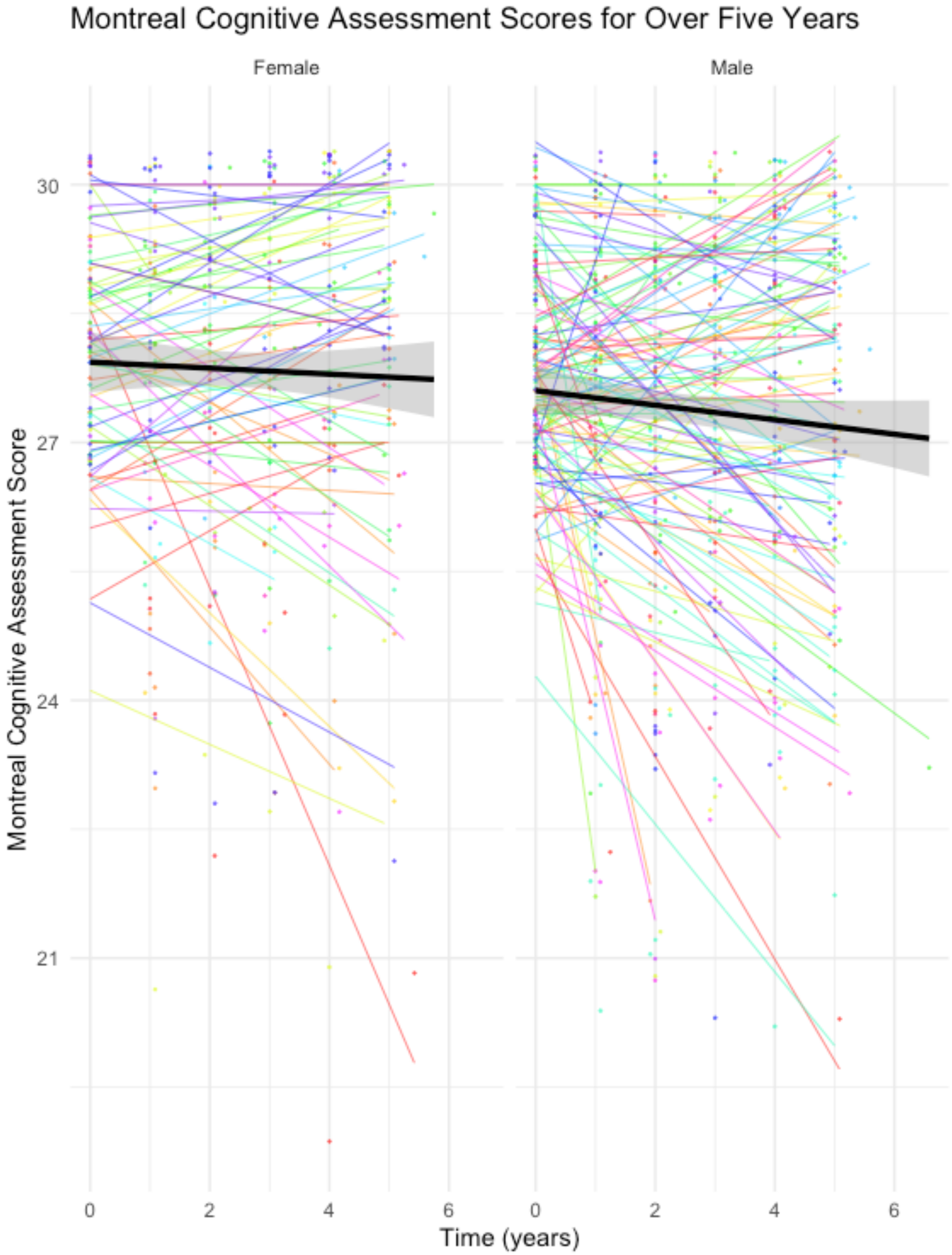

Figure 4.2 Montreal Cognitive Assessment Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in MoCA scores over time by sex with the shaded portions showing the standard error in the estimate. Note that the maximum score is 30 and some points may appear higher to show overlapping scores.

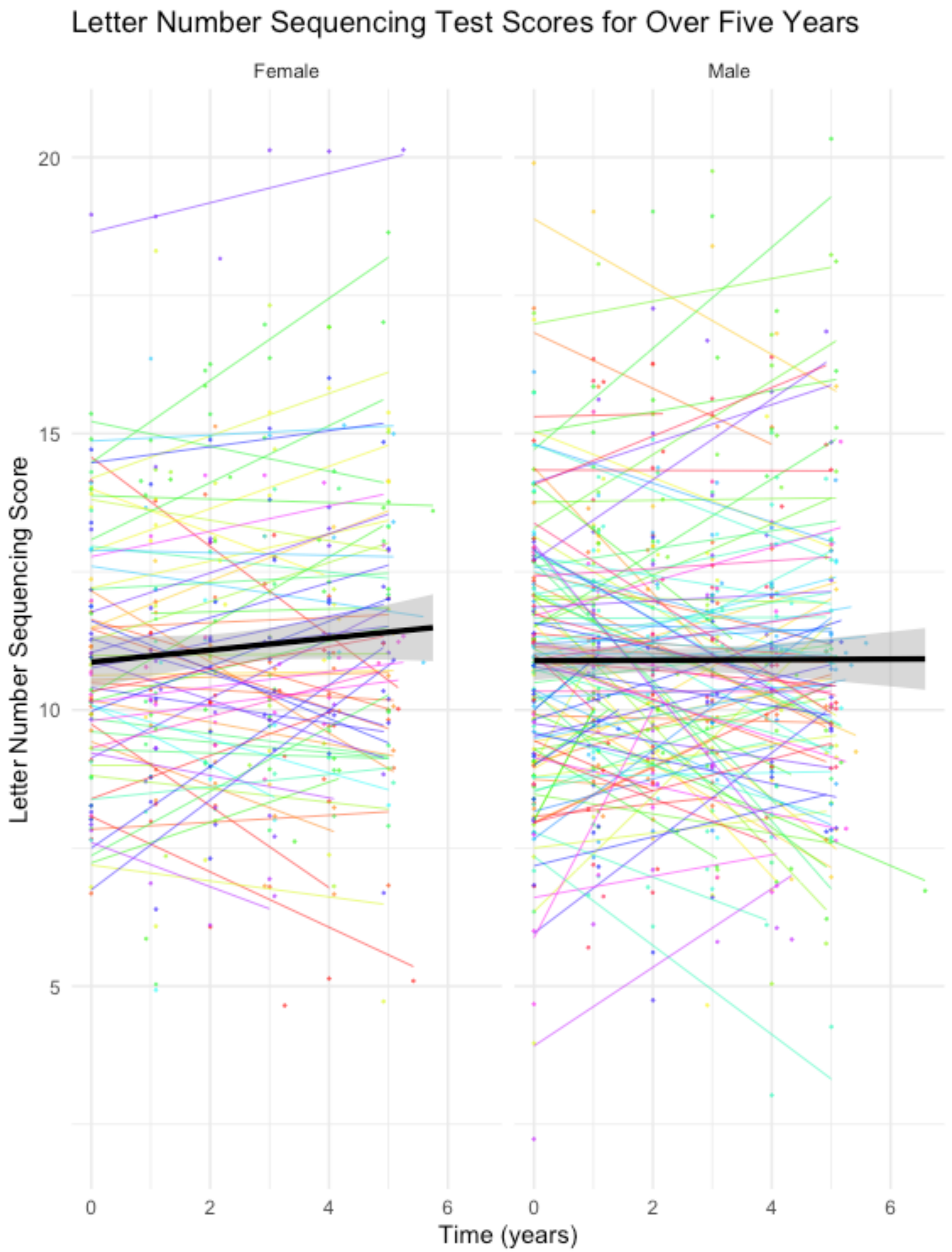

Figure 4.3 Letter Number Sequencing Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in LNS scores over time by sex with the shaded portions showing the standard error in the estimate.

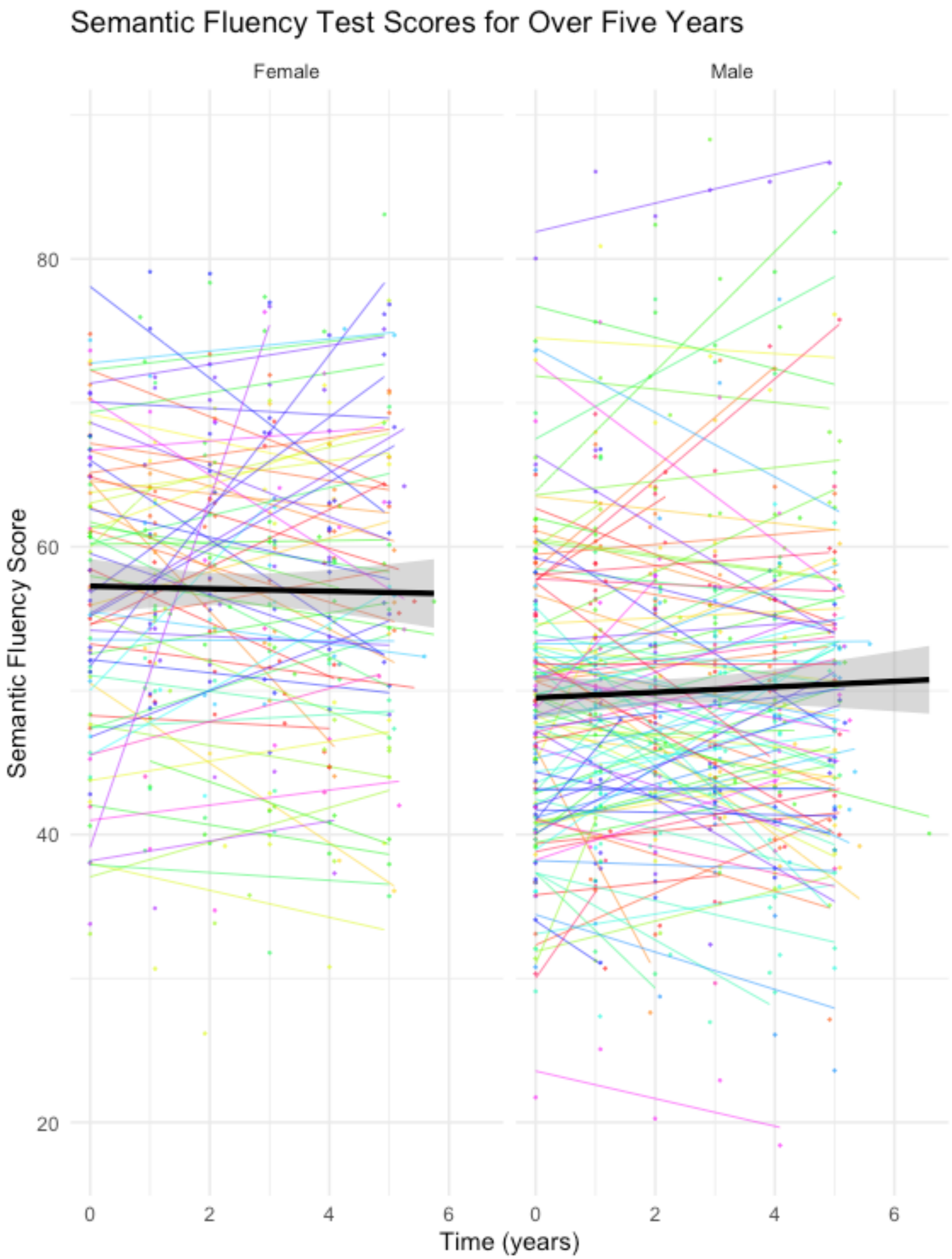

Figure 4.4 Semantic Fluency Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in SFT scores over time by sex with the shaded portions showing the standard error in the estimate.

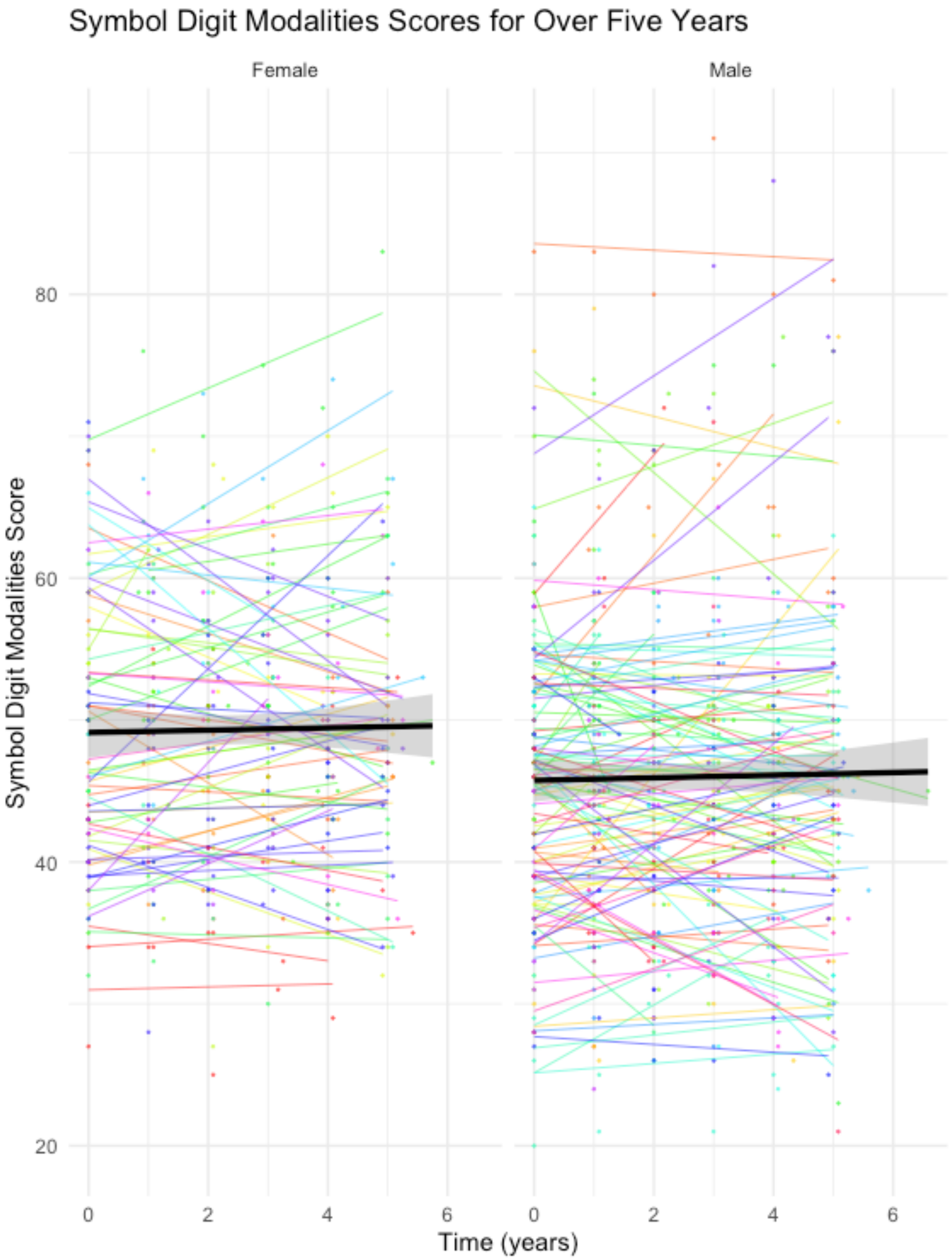

Figure 4.5 Symbol Digit Modalities Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in SDMT scores over time by sex with the shaded portions showing the standard error in the estimate.

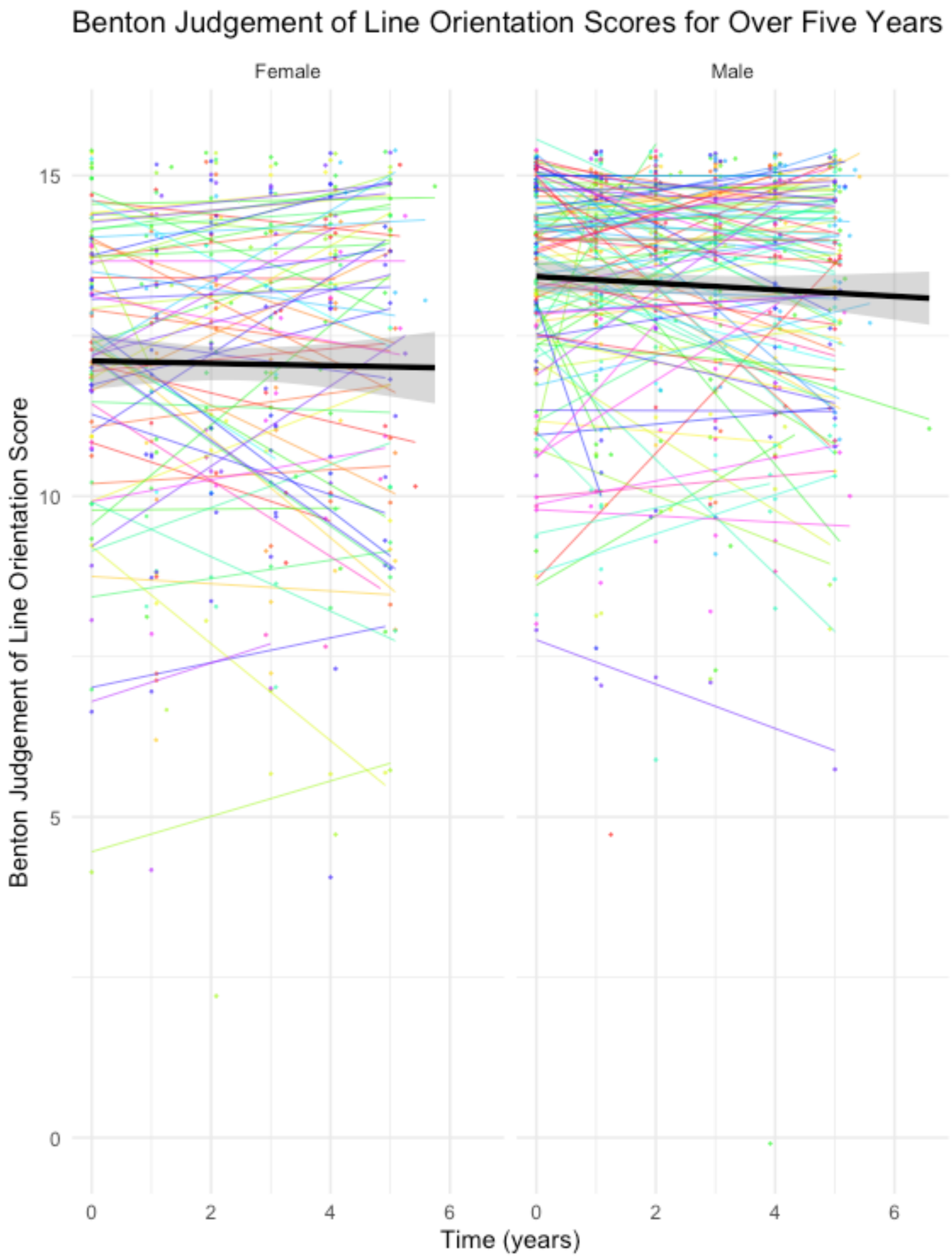

Figure 4.6 Benton Judgment of Line Orientation Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in BJLOT scores over time by sex with the shaded portions showing the standard error in the estimate. Note that the maximum score is 15 and some points may appear higher to show overlapping scores.

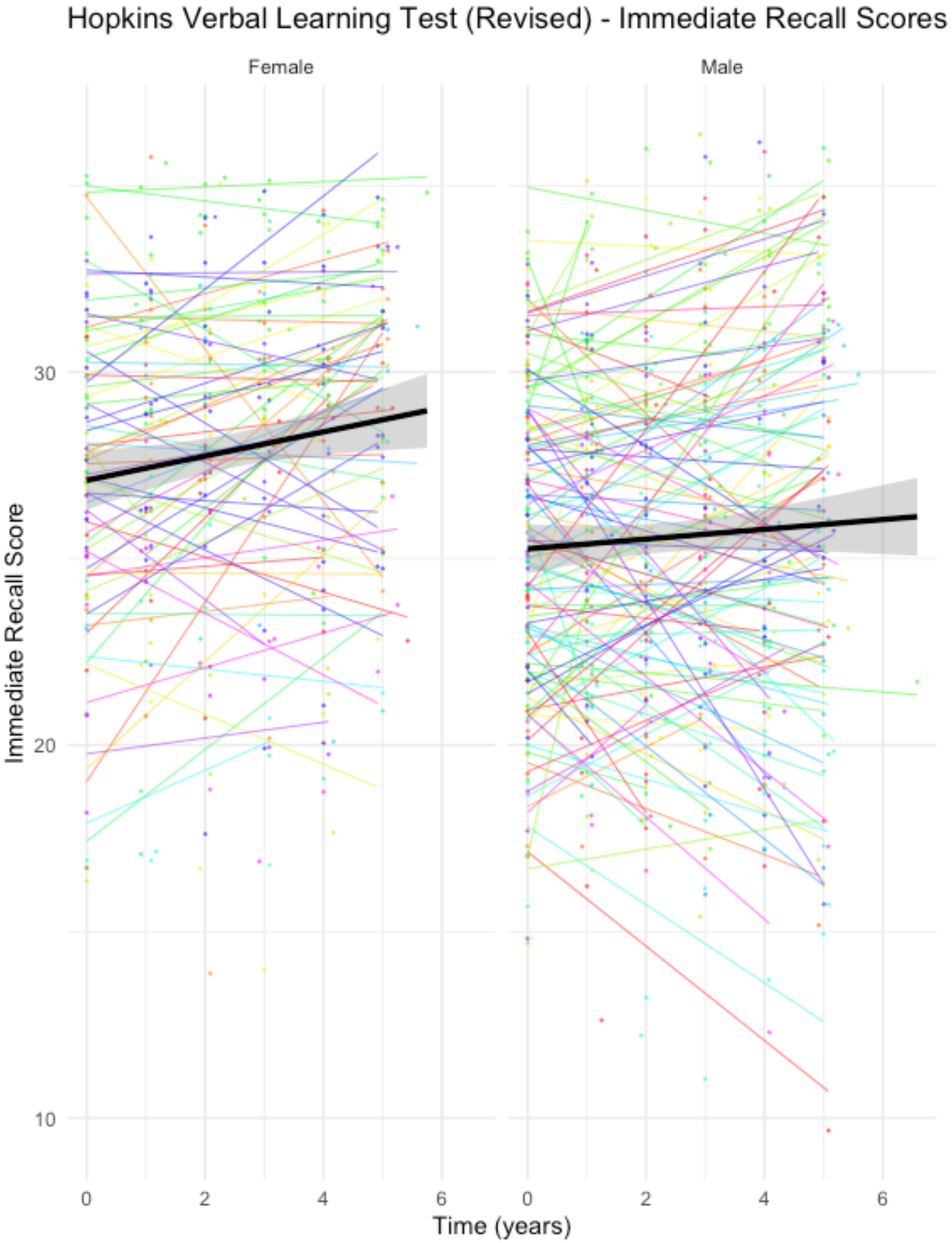

Figure 4.7 Hopkins Verbal Learning Test (Revised) Immediate Recall Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in Immediate Recall scores over time by sex with the shaded portions showing the standard error in the estimate. Note that the maximum score is 36 and some points may appear higher to show overlapping scores.

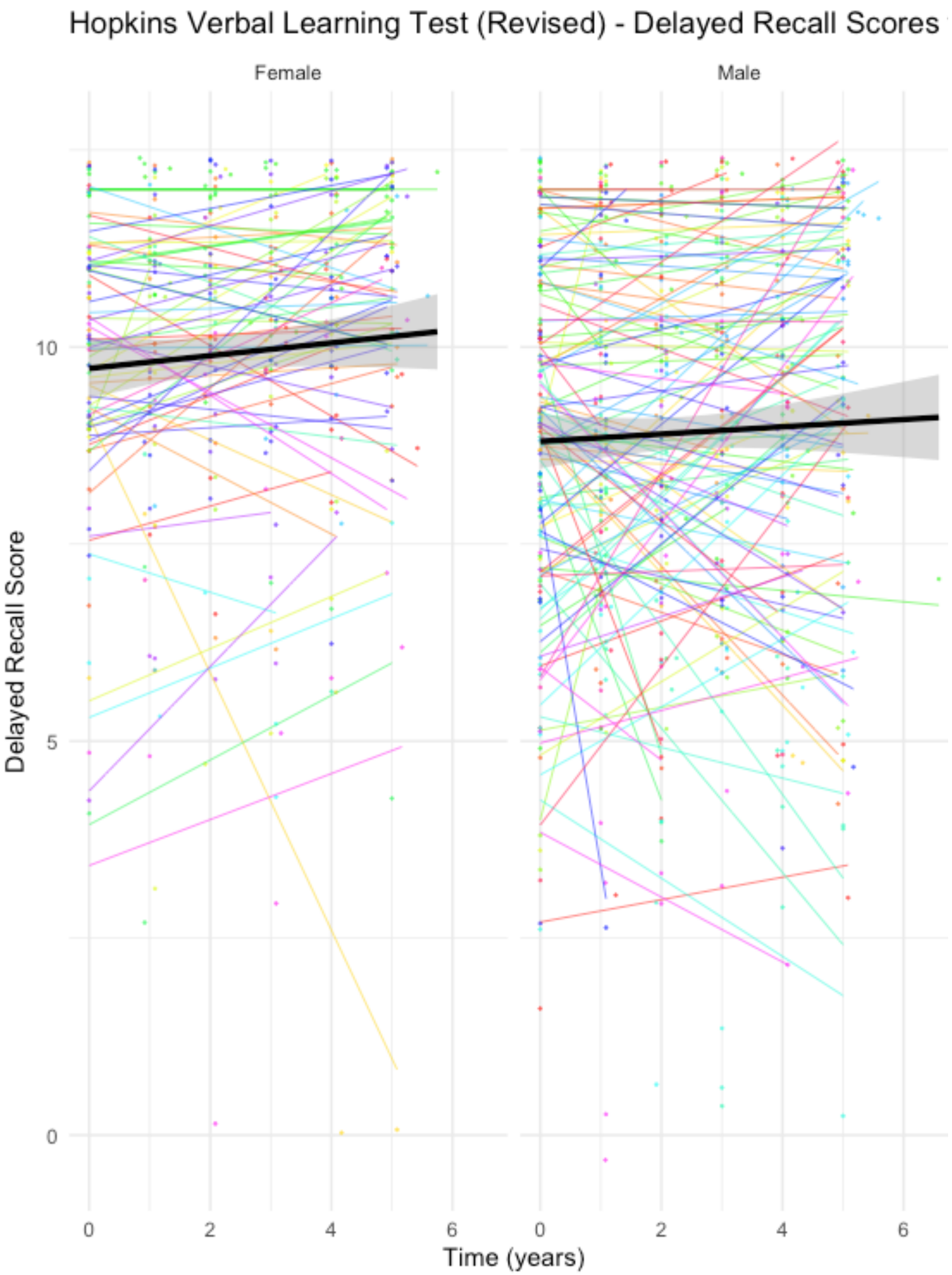

Figure 4.8. Hopkins Verbal Learning Test (Revised) Delayed Recall Scores over time.

Coloured lines show individual growth trajectories over time. Black lines show the mean change in Delayed Recall scores over time by sex with the shaded portions showing the standard error in the estimate. Note that the maximum score is 12 and some points may appear higher to show overlapping scores.

Discussion

It is important to study the healthy aging process as the population of older adults continues to rise. Although there are sex differences in life expectancy and age-related diseases, longitudinal studies investigating sex differences among healthy individuals has been limited. In this study, a secondary analysis of data sourced from the healthy control group from the PPMI was conducted. Annual cognitive changes were examined for five years (i.e., up to six time points) in 193 healthy adult males and females. A series of 2-level linear multilevel models were designed to answer three primary research questions: 1) what are the cognitive changes that occur in healthy older adults over five years (i.e., level 1), 2) what are the sex differences in cognition (i.e., level 2), and 3) what are the sex differences in rates of cognitive change over five years (i.e., level 2)? Seven neuropsychological scores were used as outcome measures and were modeled separately: MoCA, LNS, SFT, SDMT, BJLOT, HVLT-R Immediate & Delayed Recall.

Finding 1: Stability in neuropsychological scores over time

Regarding the first question of cognitive changes over time, no systematic within-person change was found on any of the neuropsychological outcome measures, except for the HVLT-R Immediate Recall. On the HVLT-R, slight increases in scores were observed over time. These results suggest overall stability among almost all neuropsychological scores in this healthy adult sample that was followed over five years.

These findings are consistent with cross-sectional and longitudinal findings of cognition in a robust sample of over 26,000 participants from the UK Biobank, which found minimal changes in cognition prior to the age of 65 (Cornelis et al., 2019). Even though many longitudinal studies of aging in older adults free of dementia have found overall decline in cognitive performance over time (Barnes et al., 2003; Gerstorf et al., 2006; McCarrey et al., 2016; Seeman et al., 2001; Singer et al., 2003), some have found overall stability in cognitive performance. For example, in the Canadian Study of Health and Aging, almost 50% of the elderly participants showed virtually no change in cognition over 10 years (McDowell et al., 2004). The present sample included the entire healthy control cohort of the PPMI, which had a wide age range at baseline from 30.63 to 83.68 years. Participants in the present sample were excluded from all time points of the healthy aging cohort by PPMI investigators if performance was more than 1.5 standard deviations below the mean on two or more neuropsychological test scores or there was PPMI investigator rating of mild cognitive impairment or dementia. Given these exclusion criteria, scores were constrained to fall within the normal range based on pre-existing normative data.

There was a slight increase in HVLT-R Immediate Recall scores over time. Significant increases in HVLT-R scores over one year were also found in a sample of over 16,000 healthy older community dwelling adults with an approximate mean age of 75 from the ASPirin in Reducing Events in the Elderly (ASPREE) study who were tested with the same form (Webb et al., 2022). Retest effects in individuals in young to middle adulthood have been demonstrated to be quite strong, such that retest effects may mask age effects or even exceed them to show positive changes over time (Salthouse, 2019; Salthouse et al., 2004). Indeed, other longitudinal studies have observed improvements in cognitive measures over time (Cornelis et al., 2019), including other verbal memory measures, such as the California Verbal Learning Test, Second Edition (Liampas et al., 2023; Lundervold et al., 2014). Alternative forms of the HVLT-R were used at each assessment, which have demonstrated acceptable reliability to the original form (Benedict et al., 1998). However, significant practice effects in serial neuropsychological testing have been found for similar verbal memory measures, such as the Rey Auditory Verbal Learning Test, even when using alternative forms (Beglinger et al., 2005).

Although virtually no systematic change in neuropsychological scores was detected in the current study, examination of the spaghetti plots (*Figures 4.2-4.8*) revealed great intraindividual variability in scores over time. Some individuals appeared to have decreases in scores over time, while others appeared to show stability or increases in scores. The ICCs for the neuropsychological outcome measures ranged from 0.58 to 0.75, confirming highly correlated scores within-persons over time and larger between-person variance in scores. This may have created a "fanning" effect that ameliorated overall systematic change over time. These findings underscore the importance of examining factors that might explain intraindividual variability in aging trajectories. For example, emerging work has strived to categorize healthy individuals into non-overlapping cognitive phenotypes based on baseline factors (e.g., demographics, psychological measures, plasma cytokines, and neuroimaging variables) and varying change trajectories within the normal range in certain cognitive domains to predict vulnerabilities in aging in healthy high functioning adults (Casaletto et al., 2019).

In the current study, several factors to explain these between-person differences in trajectories were examined. Consistent with findings in the literature (reviewed in Sanchez-Izquierdo & Fernandez-Ballesteros, 2021), higher age and lower education at baseline predicted lower performance on most neuropsychological measures, with the exception of the MoCA. There was also a statistically significant time by age interaction for the models predicting MoCA and BJLOT scores, although these estimates were exceedingly small. A primary research question examined whether there were sex differences in cognition and how rates of change may differ by sex.

Finding 2: Sex difference in neuropsychological scores at baseline

For the second question of sex differences in cognition, there were significant sex differences on several neuropsychological measures. Females scored higher on the SFT, the HVLT-R Immediate and Delayed Recall, and the SDMT, reflecting higher performance in verbal fluency, verbal learning and memory, and a multifaceted measure of attention, working memory, and information processing speed. Males scored higher on the BJLOT, reflecting higher performance in visuospatial skill. Males and females did not significantly differ on the MoCA or the LNS, reflecting similar performances in global cognitive functioning and auditory working memory.

These findings are consistent with the literature. There is a lengthy and well-documented female advantage for performance on verbal tasks, including verbal fluency and memory (Andreano & Cahill, 2009; Hyde, 2014). Recent meta-analyses support a small female advantage in verbally based episodic memory tasks (Asperholm et al., 2019; Hirnstein et al., 2023), a very small to negligible female advantage for verbal working memory (Voyer et al., 2021), and female advantage in phonemic but not semantic fluency tasks, which appear to be category dependent (Hirnstein et al., 2023). Given this literature, it is not surprising that females scored higher than males on the SFT, which was the sum of each number of category words named in one minute where there is a known advantage for females in two out of three of the categories assessed (i.e., vegetables and fruit) (Hirnstein et al., 2023). Further, females scored higher on a verbal memory measure - the

HVLT-R Immediate and Delayed Recall. Although the originally published normative data for the HVLT-R included 541 healthy adults aged 17 to 88 from the United States provided scores stratified only on age (Benedict et al., 1998; Brandt & Benedict, 2001), recently expanded normative data for older adults aged 65 included additional variables and demonstrated that females consistently scored higher than males on all components of the HVLT-R (Ryan et al., 2021). Females also scored almost three points higher than males at baseline on the SDMT. This finding is also consistent with updated normative data, where females have demonstrated higher scores than males for both the oral (Strober et al., 2020) and written version of the test (Kiely et al., 2014). Finally, males scored higher than females on the BJLOT, which is consistent with the finding of stronger performance on visuospatial ability in males compared to females (Halpern et al., 2007; meta-analysis in Voyer et al., 2017). Corrections for age and sex are typically made for the BJLOT, where males have been found to score approximately 2 points higher than females on both the full and short forms (described in Strauss et al., 2006), which was very comparable to the present findings. Similar sex differences in cognitive abilities at baseline time points have been found in several longitudinal studies of healthy aging (Aartsen et al., 2004; Cornelis et al., 2019; de Frias et al., 2006; McCarrey et al., 2016) .

The reasons for sex differences in cognitive abilities are complex and likely influenced by multiple factors. Biologically, sex hormones are thought to interact with cognition. For example, contraceptives have been linked with higher performance on verbal fluency and mental rotations tasks (Griksiene & Ruksenas, 2011). Sex hormones, such as estrogen, have also been thought to be neuroprotective and involved in neurobiological processes involved in cognition, including memory (reviewed in Gurvich et al., 2018). Hormones are also proposed to intersect with other biological processes (e.g., inflammation, vascular risk) that may further impact cognitive functioning (Szoeke et al., 2021). From a social perspective, recent research in European participants have proposed that there have been changes in the differences in cognition found between the genders. Specific cognitive abilities were found to vary systematically between men and women across birth cohorts and regions, which were associated with changes in living conditions and cognitive stimulation taking place over time (Weber et al., 2014). The authors theorized that women benefited more from these societal improvements than men, as evidenced by increased gender differences in episodic memory, decreased gender differences in numeracy, and elimination of gender differences in category fluency. Notably, sex and gender are inextricably linked, and it is likely an intersection of both biological and social factors that influence sex and gender differences in cognition.

Finding 3: No sex difference in neuropsychological scores over time

Regarding sex differences in cognitive change over time, no sex differences in rates of cognitive change were observed. That is, males and females showed similar stability in neuropsychological scores across time. This finding is consistent with many longitudinal studies of healthy aging where there were no sex differences in rates of cognitive change over time (Aartsen et al., 2004; de Frias et al., 2006; Ferreira et al., 2014; Singer et al., 2003). Sex did not determine the rate of cognitive decline in a systematic review of older adults between the ages of 60 to 80 years of age (Ferreira et al., 2014), nor were there any differences in decline in one of the most robust samples to date from the UK Biobank, which was inclusive of participants starting from age 45 (Cornelis et al., 2019). A minority of longitudinal studies have found sex differences in rates of cognitive decline, particularly in older adult samples. For example, in the Berlin Aging Study with a 13-year follow-up period, found that females had slightly steeper declines in global cognition than males (Proust-Lima et al., 2008). In the Baltimore Longitudinal Study of Aging, faster declines in males compared to females in mental status, perceptual integration and speed, and spatial ability (McCarrey et al., 2016). In the Swedish Adoption/Twin Study of Aging, there were faster declines in females on a measure of general information, and faster declines in males on a card rotation measure (Finkel et al., 2006). In a sample including middle to older adults, males showed slightly steeper declines in verbal memory compared to females (Lundervold et al., 2014).

Limitations

There are several limitations to the present study. This study was limited to the data collected by the PPMI. Consensus definitions clarify that although sex and gender may overlap for many individuals, they are distinct entities where sex describes biological factors and gender is a self-identified expression of identity. As data were only available for biological sex, the effects of gender and cognition were not specifically and separately examined in the current study. It will be important to examine both sex and gender in future research. Further, the sample of PPMI participants was largely white/Caucasian (92% of participants at baseline) and there were limited participants of other races/ethnicities, which were not well defined (i.e., "Black", "Asian", or "Other"). There is an urgent call to recruit more diverse samples in aging research as it has been found that underrepresented groups can be disproportionately excluded from participation in research studies, despite some groups having risk and burden of age-related diseases (Raman et al., 2021). Participants included in the present study met criteria for PPMI's healthy control cohort (see Appendix

A). As such, some of the participants scored below 1.5 standard deviations below the mean on a single neuropsychological measure at a single time point (e.g., MoCA score < 26). Although research consensus definitions have been created for groups along the trajectory towards age-related diseases, such as subjective cognitive decline (Jessen et al., 2014), and mild cognitive impairment (Dunne et al., 2021), there is less consensus on the definition of "normal" or "healthy" participants. Decisions about the inclusion/exclusion criteria were determined a priori. The implication is that the overall findings must be interpreted in this context, which resulted in a more restricted range of possible scores within the normal range of functioning.

Strengths/Conclusion

The overall results of the present study showed stability in cognitive functioning over five years. Although there were sex differences in neuropsychological scores on several measures consistent with the literature, the rates of change in these scores over time did not systematically differ between males and females. The current study contributes to a growing body of literature specifically examining sex similarities and differences in cognition and aging in healthy adults. A secondary analysis of already collected data from a large multi-site longitudinal study enabled an in-depth examination of changes in cognition over time. The statistical method of linear multi-level modeling poses several advantages in the examining change over time over other statistical methods such as repeated measures analysis of variance (ANOVA). In multi-level modeling, time is treated as a continuous rather than categorical variable and it can flexibly handle missing data by modeling individual growth trajectories, rather than relying on the overall group mean. This study also included comprehensive coverage of cognitive functioning with the use of seven different neuropsychological scores assessing a variety of cognitive domains. The neuropsychological measures used in the present study had strong validity and reliability, are well established in clinical practice, and had many alternate forms to minimize practice effects. The sample also included adults from age 30.62 to 83.68 at baseline, which covered middle to older adulthood. In addition to sex, age and education were examined as factors that could explain between-person differences in cognitive aging. The primary findings of this study contribute to further understanding the role of biological sex in cognitive aging.

Future directions

Future research should continue to examine the effects of sex and gender on healthy aging using the consensus definitions in the literature. As there was great intraindividual variability in healthy aging trajectories, it will be important to continue to examine factors that may predict individual changes in cognition over time.

Chapter 5: Conclusions and Future Directions

Dissertation Summary and Implication of Findings

The overall goal of the present research was to understand how the structure of the brain and cognition change during the process of healthy aging and how these changes may differ by biological sex. Three complementary studies were conducted using available neuroimaging and neuropsychological data from the healthy control group of the Parkinson's Progression Markers Initiative (PPMI). The first two studies focused the brain's white matter microstructure as measured by diffusion tensor imaging (DTI), where sex differences in white matter microstructure and changes over one year were examined. The final study focused on changes in cognition over time and as a function of sex, using neuropsychological data collected over five years. The findings of these studies contribute to our understanding of biological sex similarities and differences in brain structure and cognition in the context of healthy aging.

Studying the aging process has become more important than ever before. Globally, the population of older adults is rapidly growing due to increased life expectancy. Age is a primary risk factor for many diseases, including those that cause dementia, which create an immense burden of care on society. It is essential to understand the expected structural and functional changes that occur as part of the normal aging process. In doing so, it may be possible to identify the early changes in brain structure and function that may represent a deviation from the normal aging trajectory. Ideally, this would indicate an important point in time to implement treatment. However, treatments for many age-related diseases are not yet available. In that regard, studying healthy individuals could also provide insights into the factors that predict healthy aging, which could be harnessed in a preventative sense to promote health and longevity throughout adulthood and into the latter decades of life.

There has been a long history of sex and gender bias in biomedical research, (Plevkova et al., 2021). Inconsistent definitions of sex and gender, poor inclusion and/or unexplained exclusion of participants by sex and/or gender, and omission of sex and gender-based analyses all contribute to bias in research, which ultimately negatively impact health outcomes for people of all sexes and genders. A primary contribution of this research was to explicitly examine biological sex similarities and differences in white matter microstructure and cognition, and to determine how these variables may change differently for males and females over time. Across the three studies, the analyses were limited to biological sex, as data for gender were not available. Overall, there were mixed findings, where both sex similarities and differences were found. These findings highlight the necessity of including sex-based variables in research, as both the similarities and differences are important for understanding healthy aging.

Despite representing approximately 50% of brain tissue, limited research has examined white matter microstructure in healthy older adults and even fewer studies have explicitly examined sex similarities and differences. In Study 1, sex differences in white matter microstructure in a sample of 40 healthy older adults were examined with DTI. Using tract-based spatial statistics, white matter microstructure was compared between older adult males and females while controlling for total brain volume (TBV). No significant sex differences were found in either of the most utilized metrics of white matter microstructure - fractional anisotropy (FA) and mean diffusivity (MD). Findings from Study 1 demonstrate that older adult males and females have very similar white matter microstructure when controlling for differences in total brain volume. Other studies using similar methods (i.e., tract-based spatial statistics) have found sex differences in white matter microstructure in various structures throughout the brain (Chou et al., 2011; Kanaan et al., 2012, 2014; Menzler et al., 2011); however, direct comparison to these studies is difficult, as this research was conducted with samples that were at approximately two decades younger than the older adult sample in Study 1. The findings from Study 1 importantly contribute to our understanding of white matter microstructure in healthy older adults and add to a growing body of literature that supports the notion that brain structure in healthy males and females is similar once accounting for sex differences in gross morphology of the brain (Eliot et al., 2021).

In Study 2, longitudinal changes in white matter microstructure were examined in the same healthy older adult sample over a one-year time frame. When examining all participants, stability in FA was found, while there were significant decreases in MD in the right superior and inferior longitudinal fasciculus, the right corticospinal tract, and the right inferior fronto-occipital fasciculus. However, stability in both FA and MD were found when examining changes in white matter microstructure for males and females separately. Overall, findings from Study 2 demonstrate that healthy older adult males and females showed similar stability in white matter microstructure over a one-year period. These findings contrast with several longitudinal studies showing widespread changes in white matter microstructure with age. Most prominently observed are decreases in FA (Bender et al., 2016; Charlton et al., 2010; Lövdén et al., 2014a; Sexton et al., 2014) and increases in MD (Charlton et al., 2010; Lövdén et al., 2014b; Sexton et al., 2014). This pattern of change in white matter microstructure (i.e., decreased FA and increased MD) has also been found after a one-year period in sample of healthy control participants from the Alzheimer's Disease Neuroimaging Initiative (Mayo et al., 2017). However, other researchers have also found no differences in white matter microstructure and gender (undefined) (Ouyang et al., 2021). It is unknown if there are other unpublished studies that have also found limited sex differences due to a publication bias towards novel findings (i.e., "the file drawer problem"). The results from Studies 1 and 2 add an important contribution to the literature on sex-based similarities in white matter change over time, specifically in healthy older adults.

In Study 3, longitudinal changes in cognition as measured by several neuropsychological tests were examined in the full sample of 193 healthy control participants from the PPMI. Linear mixed modeling was used to examine within-person change over time and between-person differences in cognition for up to six annual time points. Overall stability in cognition over time was observed for all participants. Sex differences in cognition were found on several neuropsychological measures. Females scored higher than males on measures assessing verbal abilities, including verbal fluency, verbal learning and memory, and females also scored higher on a multifaceted measure of attention, working memory, and information processing speed (i.e., Semantic Fluency, Hopkins Verbal Learning Test - Revised Immediate and Delayed Recall, and Symbol Digit Modalities Test). Males scored higher than females on a measure of visuospatial ability (i.e., Benton Judgment of Line Orientation). Males and females did not significantly differ on a screener measure of global cognition or a measure of auditory working memory. There were no sex differences in rates of cognitive change over time. Findings from Study 3 demonstrated that although males and females may have differences in some cognitive domains, both sexes had similar stability in cognition over the five-year study period.

Findings from Study 3 are in line with several longitudinal studies where significant stability in cognitive scores have been found over time (Cornelis et al., 2019; McDowell et al., 2004). The areas of sex difference at baseline are consistent with reviews demonstrating a female advantage for verbal fluency and memory (Andreano & Cahill, 2009; Asperholm et al., 2019; Hirnstein et al., 2023), a female advantage on the SDMT (Kiely et al., 2014), and a male advantage for visuospatial abilities (Halpern et al., 2007; Voyer et al., 2017). A systematic review of 13 longitudinal studies also found females scored higher on measures of episodic memory, while males scored higher in visuospatial ability (Ferreira et al., 2014).

Across Studies 2 and 3, there was significant intraindividual variability in trajectories of change for both males and females. That is, some individuals demonstrated declines, some demonstrated improvements, and some demonstrated stability in white matter microstructure and cognitive functioning over time. Biological sex did not well account for these differences in individual trajectories. This is an important finding and other possible predictors that explain these varying trajectories should continue to be explored. Overall, the results of all three studies contribute to our knowledge of sex similarities and differences in healthy aging.

Strengths

There were many strengths across these three complimentary studies. The public availability of large multi-site multimodal longitudinal databases allows for cost-effective methods to study healthy adults and individuals with neurological conditions. This research was made possible by conducting secondary analyses on freely available data from the PPMI, a large multisite longitudinal study aimed at identifying biomarkers of Parkinson's disease (Marek et al., 2011b). Large multisite studies, such as the PPMI, provide a convenient and cost-effective opportunity to study healthy aging by examining data for the healthy control cohort. The present research utilized diffusion tensor imaging data as well as neuropsychological data. Many neuropsychological measures that are well established in clinical practice and those with strong psychometrics were available and utilized in this research. In the Study 3, there were no significant sex differences on a global cognitive screener, but there were significant sex differences on neuropsychological measures that assess more specific cognitive functions. These findings may have been missed without the inclusion of such a comprehensive and nuanced assessment of cognitive functioning. Using data from the PPMI enabled a comprehensive and cost-effective examination of both structural brain changes and cognition in the context of healthy aging.

Another important contribution of this research was the methodological approach to control for sex-based differences in gross morphology of the brain. Males have larger bodies and heads than females and this trend continues neuroanatomically, where on average males have 12% larger intracranial volumes, 11% larger total brain volumes, 9% larger grey matter volumes, and 13% larger white matter volumes than females (Ruigrok et al., 2014). Implications for these sex-based differences in neuroimaging research focused on white matter microstructure has rarely been addressed or discussed. However, one study has posited whether these variables should be considered a confounding or explanatory factors (Kanaan et al., 2012). A review of various intracranial volume correction methods in grey matter studies found that differences in total intracranial volume and the specific correction method used had a large impact on the estimated size of sex differences in grey matter in various structures in the brain (Sanchis-Segura et al., 2019). To date, no comparable methodological study has been conducted for investigations of white matter. In Study 1, older adult males in the sample had significantly larger total brain volumes than females, which was consistent with previous research (Eliot et al., 2021; Ruigrok et al., 2014). As sex differences in DTI metrics could theoretically be explained by differences in total brain volume, this variable was included as a covariate in the Study 1 analyses. This is one of the first studies to explicitly examine sex differences in white matter microstructure while covarying for differences in total brain volume.

Another strength of this research was its longitudinal approach. Longitudinal study design poses several advantages over cross-sectional study designs. The same participants were followed over time, which allowed for the examination of intraindividual change. Of the longitudinal studies, Study 2 followed white matter microstructural integrity in 39 healthy older adults from baseline to one year later, while Study 3 followed cognitive functioning in 193 healthy adults from age 30-83 for up to six annual time points. In Study 3, cognition was measured with six different neuropsychological measures and alternative forms for most measures were used at each time point to ameliorate practice effects (see Table 4.1).

Further, linear multi-level modeling was used to analyze the longitudinal data in Study 3. This approach has several advantages for examining change over time than over other longitudinal statistics such as repeated measures analysis of variance (ANOVA). In multi-level modeling, time was treated as a continuous rather than categorical variable. In this approach, the specific follow-up time between assessments was captured for each individual participant. This method also flexibly handled missing data by estimating individual growth trajectories using all the available data. In repeated measures ANOVA, individuals without data for all time points would have been eliminated from the study, which would have significantly reduced the sample size. The individual variation in follow-up time would not have been captured as time would have been treated as a categorical variable. And finally, change over time would have been captured by the overall group mean, rather than individual trajectories. In sum, linear mixed modelling was a strong statistical approach to understand the within and between person differences in cognition in this sample of healthy adults.

Overall, the methodology across the presented studies was strong and this research contributes to the current understanding of sex similarities and differences in healthy aging.

Limitations

There were also several limitations of the present research. This research involved secondary data analysis of the healthy control group of the PPMI, so the study samples were constrained to the available data. This posed some overall disadvantages in terms of unavailability of the variable of gender, the definition of a "healthy" participant, and the overall lack of ethnic diversity within the sample. There are also some limitations to the methodology used to analyze the diffusion tensor imaging data, which could potentially impact the findings of Study 1 and 2. These limitations are important for understanding the generalizability the findings and informing future research.

This research was limited to the study of biological sex differences as data for gender were not available. Although consensus definitions of sex and gender have been published (see the summary in Table 1.1), these terms continue to be used interchangeably in the literature. Indeed, the PPMI database published this variable as "gender at birth", which corresponds with the consensus-based definition of sex encompassing biological and physical attributes of a person (Marek et al., 2011a). There is a bidirectional entanglement of sex and gender, which makes it challenging to separate these variables (Eliot et al., 2021). However, gender aligns with sex assigned at birth for most people. In 2021, Canada became the first country in the world to collect and publish data on gender diversity via a national census. These data showed 0.33% or approximately 1 in 300 Canadians aged 15 and older reported themselves to be transgender or non-binary (Statistics Canada, 2022). This suggests that the results from this study may be applicable for male or female cis-gendered individuals (i.e., people whose gender aligns with their sex assigned at birth).

Participants included in these studies met criteria for PPMI's healthy control participants (see Appendix A). Although research consensus definitions have been created for groups along the trajectory towards age-related diseases, such as subjective cognitive decline (Jessen et al., 2014), and mild cognitive impairment (Dunne et al., 2021), there is less consensus on the definition of "normal" or "healthy" participants. In the present research, healthy control participants from the PPMI scored 26 or greater on a screening measure of cognition (i.e., the MoCA). On subsequent assessments, they had to score more than 1.5 deviations below the mean on two or more neuropsychological tests or have a PPMI investigator rating of mild cognitive impairment or dementia. This is why some individuals scored below 26 on the MoCA after the baseline time point in Study 2 and 3, but were not excluded. Importantly, this performance-based criteria omits information about self-reported concern with cognitive functioning, which is likely a meaningful indicator of early brain changes. In our past research, individuals who perform normally on neuropsychological measures, but had concerns about their cognitive abilities had lower white matter integrity relative to those who also performed normally on neuropsychological measures, but had no concerns about their cognitive ability (Ohlhauser et al., 2018). Our past research has also shown detectable brain differences in individuals with subjective cognitive decline relative to healthy controls without such concerns (Parker et al., 2022).

Nonetheless, all participants met the criteria for healthy control participants outline in Appendix A for every time point for which data was available. This research tracked change among individuals who remained healthy and did not convert to mild cognitive impairment or dementia within the study period. In this sense, it is not surprising that little variation in cognition was seen over time, as the criteria for healthy control participants in the PPMI generally constrained the possible variation in scores to be at or above a normative cut-off score of 1.5 standard deviations below the mean.

Another limitation of using data from the PPMI was the lack of diversity of within this healthy control cohort. The data was collected in the United States with participants who were mostly Caucasian. Limited research has examined ethnicity or social determinants in brain health. The Healthy Aging in Neighbourhoods of Diversity Across the Life Span or HANDLS, is a multidisciplinary, community-based, prospective longitudinal epidemiologic study examining the influences of race and socioeconomic status on the development of age-related health disparities among socioeconomically diverse African

Americans and white participants in Baltimore, USA. In what is thought to be one of the first studies to independently examine the interactive relationship between socioeconomic status, race, and age with white matter integrity, those with lower socioeconomic status, higher age, and male sex, had lower FA (reflecting poorer white matter integrity), while there were no significant differences in white matter integrity based on race (African American or White) (Shaked et al., 2019). Based on these findings, ethnicity may be associated with other factors that impact health outcomes, such as socioeconomic status. Nonetheless, as such limited research has been conducted in this area, it will be important to strive to include people of diverse ethnic and socioeconomic backgrounds into healthy aging research.

Regarding the neuroimaging studies (Study 1 and 2), there are some limitations to the methodology used to analyze the diffusion tensor imaging data. Tract-Based Spatial Statistics (TBSS) is an approach to objectively estimate the overall white matter tracts within the brain that are common to study subjects, which can then be compared statistically (Smith et al., 2006). Although the recommended processing pipeline was followed, this method has several known limitations. These include possible anatomic inaccuracies and bias in the skeleton projection step, which can potentially cause problems when analyzing small tubular structures due to inter-subject differences in its location and crossing or kissing fibres that can alter DTI metrics (Bach et al., 2014). Further, the interpretation of DTI metrics is not simple as “increased MD equals neurodegeneration” as increased in MD can also be attributed to other factors, such as changes in partial volume effects due to characteristics in fiber bundles such as thickness, orientation, and curvature (Vos et al., 2011). Despite these limitations, these results contribute to our understanding of sex-based similarities and differences in white matter microstructure and cognition in healthy adults.

Future Directions

Future research should continue to examine sex and gender similarities and differences in healthy aging. Clear definitions of sex and gender and the inclusion of sex and gender-based analyses are needed. Strategies for integrating sex and gender into neurodegeneration research in Canada have been published, which encompasses research in healthy aging. Six components have been proposed, including: executive oversight, training, research collaboration, progress report assessment, results dissemination, and ongoing manuscript review (Tierney et al., 2017). Other guidelines for research design in basic, clinical, and population studies have also been published to drive scientific and therapeutic discovery for across the spectrum of sex and gender (Rich-Edwards et al., 2018).

Further, little is known about how the aging process may differ for sex and gender diverse populations. Including those who are intersex, transgender, and/or non-binary individuals in future research is important. Data from the Canadian census suggests that there is a significant generational gap in reporting of transgender and non-binary people, where reporting rates were three to seven times higher in Generation Z (people born between 1997 and 2006), compared to older generational cohorts (Statistics Canada, 2022). It is likely that this research may become more prevalent with time as awareness and acceptance of gender diverse populations improves. Much of the present research is limited to a binary measurement of biological sex as either male or female, which limits the applicability of these results for sexual and gender diverse minorities. Importantly, comparisons of various sexual and gender diverse populations could also provide insights into biological and sociological determinants of health. For example, regarding biological sex, females with complete androgen insensitivity syndrome (i.e., those who lack androgen action in the presence of a 46,XY karyotype) had white matter microstructural integrity more closely resembling 46,XX females (i.e., cisgendered females) than 46,XY males (i.e., cisgendered males), suggesting that sex hormones played a greater role in the expression of white matter integrity than sex chromosomes (van Hemmen et al., 2017). Further research is needed to understand the role of sex and gender in white matter integrity and healthy aging in general.

Future research should also carefully examine definitions of "normal" or "healthy" individuals in the context of cognition. Notably, brain differences have been found between individuals who perform normally on cognitive measures with subjective cognitive complaints compared to individuals without such complaints (Ohlhauser et al., 2018; Parker et al., 2022). These findings suggest that subjective cognitive concerns are an important variable to consider when investigating healthy individuals and should be included as a routine practice when collecting data on healthy cohorts.

For future research, it is also important to recognize that there are significant differences in performances on the same neuropsychological tests among different normative samples, even within samples thought to be quite culturally similar. For example, Canadians score higher than Americans all of the Wechsler intelligence scales throughout the lifespan, which necessitated the development of Canadian based normative data (Miller et al., 2015). Further, the MoCA was originally developed using a sample of

90 normal control participants from Canada (mean age=72.84, SD=7.03, education=13.33, SD=3.40) (Nasreddine et al., 2005). In clinical practice the original cut-off score of 26 is well established, which is based on optimal sensitivity and specificity for detecting mild cognitive impairment and Alzheimer's disease. However, in a normative study using an ethnically diverse sample from the United States, the majority of MoCA scores fell below 26 (Rossetti et al., 2011), but were reflective of normal cognition. It is unlikely these factors drastically impacted the current study as the sample was well matched to the normative data available; however, future research should carefully examine and use appropriate normative data for their sample.

The findings from the longitudinal analyses revealed significant intraindividual variability in both white matter microstructural changes (Study 2) and cognitive changes over time (Study 3). The factors that explain these variable trajectories should continue to be explored. Although sex predicted differences in initial cognitive abilities, sex provided little predictive power to explain changes in white matter microstructure or cognition over time. These results revealed great intraindividual variability in cognitive change over time, suggesting that factors other than sex could be identified to better explain this individual variability. That is, it will also be important to identify whether certain trajectories among healthy individuals might predict future improvement, stability, or decline in brain structure and cognitive functioning. For example, emerging work has strived to categorize healthy individuals into non-overlapping cognitive phenotypes based on baseline factors (e.g., demographics, psychological measures, plasma cytokines, and neuroimaging variables) and varying change trajectories within the normal range in certain cognitive domains to predict vulnerabilities in aging in healthy high functioning adults (Casaletto et al., 2019). Indeed, this type of research aligns well with the Scaffolding Theory of Aging and Cognition Revised (STAC-r) that posit that neuronal enrichment and depletion factors throughout the life course are predictors for structural and functional brain changes (Reuter-Lorenz & Park, 2014). These neuronal enrichment and depletion factors include modifiable activities (e.g., exercise, nutrition) that could potentially improve brain structure and function, and positively benefit everyday functioning. These factors should continue to be explored in larger multisite longitudinal studies that include measurements of brain structure and function in the context of healthy aging.

Conclusion

The overall goal of this research was to better understand brain structure and function in healthy adults and how these changes may vary by biological sex and over time. This manuscript-based dissertation contained three complementary studies, which explicitly examined biological sex differences in white matter microstructure, cognition, and the changes in these factors over time in a sample of healthy adults who were healthy control participants from the PPMI. The results of Study 1 demonstrated that after accounting for sex difference in total brain volume, healthy older adult males and females have similar white matter microstructure. Studies 2 and 3 examined longitudinal changes in white matter microstructure and cognition, respectively. Both studies demonstrated overall stability in microstructure and cognitive changes over time. Biological sex accounted for baseline differences in specific cognitive domains in areas consistent with the literature. That is, females performed higher on measures of verbal fluency and memory, while males performed higher on a measure of visuospatial ability. However, both males and females showed stability in cognition over time. These three studies contribute to our knowledge of biological sex similarities and differences in brain structure and cognition and their changes over time in healthy adults. This work ultimately aids the overarching and long-term goals of aging research, which are to prevent disease and promote health and well-being throughout the lifespan and into the latter decades of life.

References

Aartsen, M. J., Martin, M., & Zimprich, D. (2004). Gender Differences in Level and Change in Cognitive Functioning. *Gerontology*, *50*(1), 35–38. https://doi.org/10.1159/000074387

Andersson, J., Jenkinson, M., & Smith, S. (2007). *Non-linear registration aka Spatial normalisation*. FMRIB Technial Report TR07JA2.

Andersson, J. L. R., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078. https://doi.org/10.1016/j.neuroimage.2015.10.019

Andreano, J. M., & Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning & Memory*, *16*(4), 248–266. https://doi.org/10.1101/lm.918309

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*, 1745506518817995. https://doi.org/10.1177/1745506518817995

Ardekani, B. A., Figarsky, K., & Sidtis, J. J. (2013). Sexual Dimorphism in the Human Corpus Callosum: An MRI Study Using the OASIS Brain Database. *Cerebral Cortex*, *23*(10), 2514–2520. https://doi.org/10.1093/cercor/bhs253

Asperholm, M., Högman, N., Rafi, J., & Herlitz, A. (2019). What did you do yesterday? A meta-analysis of sex differences in episodic memory. *Psychological Bulletin*, *145*(8), 785–821. https://doi.org/10.1037/bul0000197

Bach, M., Laun, F. B., Leemans, A., Tax, C. M. W., Biessels, G. J., Stieltjes, B., & Maier-Hein, K. H. (2014). Methodological considerations on tract-based spatial statistics (TBSS). *NeuroImage*, *100*, 358–369. https://doi.org/10.1016/j.neuroimage.2014.06.021

Banister, E. M., Coen, S., Canadian Institutes of Health Research, & Institute of Gender and Health (Canada). (2012). *What a difference sex and gender make: A gender, sex and health research casebook*. Canadian Institutes of Health Research. https://www.deslibris.ca/ID/231706

Barel, E., & Tzischinsky, O. (2018). Age and Sex Differences in Verbal and Visuospatial Abilities. *Advances in Cognitive Psychology*, *14*(2), 51–61. https://doi.org/10.5709/acp-0238-x

Barford, A., Dorling, D., Smith, G. D., & Shaw, M. (2006). Life expectancy: Women now on top everywhere. *BMJ*, *332*(7545), 808. https://doi.org/10.1136/bmj.332.7545.808

Barnes, L. L., Wilson, R. S., Schneider, J. A., Bienias, J. L. S., Evans, D. A., & Bennett, D. A. (2003). Gender, cognitive decline, and risk of AD in older persons. *Neurology*, *60*(11), 1777–1781. https://doi.org/10.1212/01.WNL.0000065892.67099.2A

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*(1), 259–267.

Bates, D., Maechler, M., Bolker [aut, B., cre, Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., Fox, J., Bauer, A., & simulate.formula), P. N. K. (shared copyright on. (2022). *lme4: Linear*

Mixed-Effects Models using "Eigen" and S4 (1.1-30). https://CRAN.R-project.org/package=lme4

Beglinger, L., Gaydos, B., Tangphaodaniels, O., Duff, K., Kareken, D., Crawford, J., Fastenau, P., & Siemers, E. (2005). Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*, *20*(4), 517–529. https://doi.org/10.1016/j.acn.2004.12.003

Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. A. M., Boulby, P. A., Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J. M., & Matthews, P. M. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *NATURE NEUROSCIENCE*, *6*(7), 8.

Bender, A. R., Völkle, M. C., & Raz, N. (2016). Differential aging of cerebral white matter in middle-aged and older adults: A seven-year follow-up. *NeuroImage*, *125*, 74–83. https://doi.org/10.1016/j.neuroimage.2015.10.030

Benedict, R. H. B., Schretlen, D., Groninger, L., & Brandt, J. (1998). Hopkins Verbal Learning Test – Revised: Normative Data and Analysis of Inter-Form and Test-Retest Reliability. *The Clinical Neuropsychologist*, *12*(1), 43–55. https://doi.org/10.1076/clin.12.1.43.1726

Benton, A. L. (Ed.). (1994). *Contributions to neuropsychological assessment: A clinical manual* (2nd ed). Oxford University Press.

Bethlehem, R. a. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., … Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), Article 7906. https://doi.org/10.1038/s41586-022-04554-y

Bishop, K. M., & Wahlsten, D. (1997). Sex Differences in the Human Corpus Callosum: Myth or Reality? *Neuroscience & Biobehavioral Reviews*, *21*(5), 581–601. https://doi.org/10.1016/S0149-7634(96)00049-8

Brandt, J., & Benedict, R. H. B. (2001). *Hopkins Verbal Learning Test-Revised*. PAR.

Casaletto, K. B., Elahi, F. M., Staffaroni, A. M., Walters, S., Contreras, W. R., Wolf, A., Dubal, D., Miller, B., Yaffe, K., & Kramer, J. H. (2019). Cognitive aging is not created equally: Differentiating unique cognitive phenotypes in "normal" adults. *Neurobiology of Aging*, *77*, 13–19. https://doi.org/10.1016/j.neurobiolaging.2019.01.007

Charlton, R. A., Schiavone, F., Barrick, T. R., Morris, R. G., & Markus, H. S. (2010). Diffusion tensor imaging detects age related white matter change over a 2 year follow-up which is associated with working memory decline. *Journal of Neurology, Neurosurgery & Psychiatry*, *81*(1), 13–19. https://doi.org/10.1136/jnnp.2008.167288

Chou, K.-H., Cheng, Y., Chen, I.-Y., Lin, C.-P., & Chu, W.-C. (2011). Sex-linked white matter microstructure of the social and analytic brain. *NeuroImage*, *54*(1), 725–733. https://doi.org/10.1016/j.neuroimage.2010.07.010

Clayton, J. A., & Tannenbaum, C. (2016). Reporting Sex, Gender, or Both in Clinical Research? *JAMA*, *316*(18), 1863–1864. https://doi.org/10.1001/jama.2016.16405

Cornelis, M. C., Wang, Y., Holland, T., Agarwal, P., Weintraub, S., & Morris, M. C. (2019). Age and cognitive decline in the UK Biobank. *PLOS ONE*, *14*(3), e0213948. https://doi.org/10.1371/journal.pone.0213948

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., Wardlaw, J. M., Gale, C. R., Bastin, M. E., & Deary, I. J. (2016). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms13629

David, S. P., Naudet, F., Laude, J., Radua, J., Fusar-Poli, P., Chu, I., Stefanick, M. L., & Ioannidis, J. P. A. (2018). Potential Reporting Bias in Neuroimaging Studies of Sex Differences. *Scientific Reports*, *8*(1), 6082. https://doi.org/10.1038/s41598-018-23976-1

Day, S., Mason, R., Lagosky, S., & Rochon, P. A. (2016). Integrating and evaluating sex and gender in health research. *Health Research Policy and Systems*, *14*(1). https://doi.org/10.1186/s12961-016-0147-7

de Frias, C. M., Nilsson, L.-G., & Herlitz, A. (2006). Sex Differences in Cognition are Stable Over a 10-Year Period in Adulthood and Old Age. *Aging, Neuropsychology, and Cognition*, *13*(3–4), 574–587. https://doi.org/10.1080/13825580600678418

Driesen, N. R., & Raz, N. (1995). The influence of sex, age, and handedness on corpus callosum morphology: A meta-analysis. *Psychobiology*, *23*(3), 240–247. https://doi.org/10.3758/BF03332028

Dunne, R. A., Aarsland, D., O'Brien, J. T., Ballard, C., Banerjee, S., Fox, N. C., Isaacs, J. D., Underwood, B. R., Perry, R. J., Chan, D., Dening, T., Thomas, A. J., Schryer, J., Jones, A.-M., Evans, A. R., Alessi, C., Coulthard, E. J., Pickett, J., Elton, P., … Burns, A. (2021). Mild Cognitive Impairment: The Manchester consensus. *Age and Ageing*, *50*(1), 72–80. https://doi.org/10.1093/ageing/afaa228

Eliot, L., Ahmed, A., Khan, H., & Patel, J. (2021). Dump the “dimorphism”: Comprehensive synthesis of human brain studies reveals few male-female differences beyond size. *Neuroscience & Biobehavioral Reviews*, *125*, 667–697. https://doi.org/10.1016/j.neubiorev.2021.02.026

Ferreira, L., Santos-Galduróz, R. F., Ferri, C. P., & Galduróz, J. C. F. (2014). Rate of cognitive decline in relation to sex after 60 years-of-age: A systematic review. *Geriatrics & Gerontology International*, *14*(1), 23–31. https://doi.org/10.1111/ggi.12093

Figley, C. R., Uddin, M. N., Wong, K., Kornelsen, J., Puig, J., & Figley, T. D. (2022). Potential Pitfalls of Using Fractional Anisotropy, Axial Diffusivity, and Radial Diffusivity as Biomarkers of Cerebral White Matter Microstructure. *Frontiers in Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnins.2021.799576

Finkel, D., Reynolds, C. A., Berg, S., & Pedersen, N. L. (2006). Surprising lack of sex differences in normal cognitive aging in twins. *International Journal of Aging & Human Development*, *62*(4), 335–357. https://doi.org/10.2190/C39X-9QHY-49DM-X9GJ

Galea, L. A. M., Frick, K. M., Hampson, E., Sohrabji, F., & Choleris, E. (2017). Why estrogens matter for behavior and brain health. *Neuroscience and Biobehavioral Reviews*, *76*(Pt B), 363–379. https://doi.org/10.1016/j.neubiorev.2016.03.024

Gao, Z., Chen, Z., Sun, A., & Deng, X. (2019). Gender differences in cardiovascular disease. *Medicine in Novel Technology and Devices*, *4*, 100025. https://doi.org/10.1016/j.medntd.2019.100025

Gerstorf, D., Herlitz, A., & Smith, J. (2006). Stability of Sex Differences in Cognition in Advanced Old Age: The Role of Education and Attrition. *The Journals of Gerontology: Series B*, *61*(4), P245–P249. https://doi.org/10.1093/geronb/61.4.P245

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders*. Lea & Febiger.

Government of Canada, C. I. of H. R. (2014, January 10). *What is gender? What is sex? - CIHR*. https://cihr-irsc.gc.ca/e/48642.html

Griksiene, R., & Ruksenas, O. (2011). Effects of hormonal contraceptives on mental rotation and verbal fluency. *Psychoneuroendocrinology*, *36*(8), 1239–1248. https://doi.org/10.1016/j.psyneuen.2011.03.001

Grundy, E., & Sloggett, A. (2003). Health inequalities in the older population: The role of personal capital, social resources and socio-economic circumstances. *Social Science & Medicine (1982)*, *56*(5), 935–947. https://doi.org/10.1016/s0277-9536(02)00093-x

Gubbels Bupp, M. R., Potluri, T., Fink, A. L., & Klein, S. L. (2018). The Confluence of Sex Hormones and Aging on Immunity. *Frontiers in Immunology*, *9*, 1269. https://doi.org/10.3389/fimmu.2018.01269

Gurvich, C., Hoy, K., Thomas, N., & Kulkarni, J. (2018). Sex Differences and the Influence of Sex Hormones on Cognition through Adulthood and the Aging Process. *Brain Sciences*, *8*(9), 163. https://doi.org/10.3390/brainsci8090163

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The Science of Sex Differences in Science and Mathematics. *Psychological Science in the Public Interest*, *8*(1), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x

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

Hirnstein, M., Stuebs, J., Moè, A., & Hausmann, M. (2023). Sex/Gender Differences in Verbal Fluency and Verbal-Episodic Memory: A Meta-Analysis. *Perspectives on Psychological Science*, *18*(1), 67–90. https://doi.org/10.1177/17456916221082116

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

Hudson, J. I., Hiripi, E., Pope, H. G., & Kessler, R. C. (2007). The Prevalence and Correlates of Eating Disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*, *61*(3), 348–358. https://doi.org/10.1016/j.biopsych.2006.03.040

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Hyde, J. S. (2014). Gender Similarities and Differences. *Annual Review of Psychology*, *65*(1), 373–398. https://doi.org/10.1146/annurev-psych-010213-115057

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. https://doi.org/10.1037/amp0000307

Inano, S., Takao, H., Hayashi, N., Abe, O., & Ohtomo, K. (2011). Effects of Age and Gender on White Matter Integrity. *American Journal of Neuroradiology*, *32*(11), 2103–2109. https://doi.org/10.3174/ajnr.A2785

Institute of Medicine (U.S.), Blazer, D. G., Yaffe, K., & Liverman, C. T. (Eds.). (2015). *Cognitive aging: Progress in understanding and opportunities for action*. The National Academies Press.

Jahanshad, N., & Thompson, P. M. (2017). Multimodal neuroimaging of male and female brain structure in health and disease across the life span. *Journal of Neuroscience Research*, *95*(1–2), 371–379. https://doi.org/10.1002/jnr.23919

Janicki, S. C., & Schupf, N. (2010). Hormonal Influences on Cognition and Risk for Alzheimer Disease. *Current Neurology and Neuroscience Reports*, *10*(5), 359–366. https://doi.org/10.1007/s11910-010-0122-6

Jellinger, K. A. (2003). Age-associated prevalence and risk factors of Lewy body pathology in a general population. *Acta Neuropathologica*, *106*(4), 383–384. https://doi.org/10.1007/s00401-003-0751-9

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jessen, F., Amariglio, R. E., van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., van der Flier, W. M., Glodzik, L., van Harten, A. C., de Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., … Wagner, M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer’s disease. *Alzheimer’s & Dementia : The Journal of the Alzheimer’s Association*, *10*(6), 844–852. https://doi.org/10.1016/j.jalz.2014.01.001

Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., Smith, S. M., Higham, D. J., & Matthews, P. M. (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(36), 13335–13340. https://doi.org/10.1073/pnas.0403743101

Johnson, J., Sharman, Z., Vissandjée, B., & Stewart, D. E. (2014). Does a Change in Health Research Funding Policy Related to the Integration of Sex and Gender Have an Impact? *PLoS ONE*, *9*(6), e99900. https://doi.org/10.1371/journal.pone.0099900

Jordan-Young, R., & Rumiati, R. I. (2012). Hardwired for Sexism? Approaches to Sex/Gender in Neuroscience. *Neuroethics*, *5*(3), 305–315. https://doi.org/10.1007/s12152-011-9134-4

Kaiser, A. (2012). Re-conceptualizing “sex” and “gender” in the human brain. *Zeitschrift Für Psychologie*, *220*(2), 130–136. https://doi.org/10.1027/2151-2604/a000104

Kanaan, R. A., Allin, M., Picchioni, M., Barker, G. J., Daly, E., Shergill, S. S., Woolley, J., & McGuire, P. K. (2012). Gender differences in white matter microstructure. *PLoS ONE*, *7*(6), e38272. https://doi.org/10.1371/journal.pone.0038272

Kanaan, R. A., Chaddock, C., Allin, M., Picchioni, M. M., Daly, E., Shergill, S. S., & McGuire, P. K. (2014). Gender Influence on White Matter Microstructure: A Tract-Based Spatial Statistics Analysis. *PLoS ONE*, *9*(3), e91109. https://doi.org/10.1371/journal.pone.0091109

Kiely, K. M., Butterworth, P., Watson, N., & Wooden, M. (2014). The Symbol Digit Modalities Test: Normative Data from a Large Nationally Representative Sample of Australians. *Archives of Clinical Neuropsychology*, *29*(8), 767–775. https://doi.org/10.1093/arclin/acu055

Kochunov, P., Williamson, D. E., Lancaster, J., Fox, P., Cornell, J., Blangero, J., & Glahn, D. C. (2012). Fractional anisotropy of water diffusion in cerebral white matter across the lifespan. *Neurobiology of Aging*, *33*(1), 9–20. https://doi.org/10.1016/j.neurobiolaging.2010.01.014

Kopecek, M., Bezdicek, O., Sulc, Z., Lukavsky, J., & Stepankova, H. (2017). Montreal Cognitive Assessment and Mini-Mental State Examination reliable change indices in healthy older adults. *International Journal of Geriatric Psychiatry*, *32*(8), 868–875. https://doi.org/10.1002/gps.4539

Levinsson, A., de Denus, S., Sandoval, J., Lemieux Perreault, L.-P., Rouleau, J., Tardif, J.-C., Hussin, J., & Dubé, M.-P. (2022). Construction of a femininity score in the UK Biobank and its association with angina diagnosis prior to myocardial infarction. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-05713-x

Liampas, I., Folia, V., Ntanasi, E., Yannakoulia, M., Sakka, P., Hadjigeorgiou, G., Scarmeas, N., Dardiotis, E., & Kosmidis, M. H. (2023). Longitudinal episodic memory trajectories in older adults with normal cognition. *The Clinical Neuropsychologist*, *37*(2), 304–321. https://doi.org/10.1080/13854046.2022.2059011

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lövdén, M., Köhncke, Y., Laukka, E. J., Kalpouzos, G., Salami, A., Li, T.-Q., Fratiglioni, L., & Bäckman, L. (2014a). Changes in perceptual speed and white matter microstructure in the corticospinal tract are associated in very old age. *NeuroImage*, *102*, 520–530. https://doi.org/10.1016/j.neuroimage.2014.08.020

Lövdén, M., Köhncke, Y., Laukka, E. J., Kalpouzos, G., Salami, A., Li, T.-Q., Fratiglioni, L., & Bäckman, L. (2014b). Changes in perceptual speed and white matter microstructure in the corticospinal tract are associated in very old age. *NeuroImage*, *102*, 520–530. https://doi.org/10.1016/j.neuroimage.2014.08.020

Lovden, M., Ronnlund, M., Wahlin, A., Backman, L., Nyberg, L., & Nilsson, L.-G. (2004). The Extent of Stability and Change in Episodic and Semantic Memory in Old Age: Demographic Predictors of Level and Change. *The Journals of*

Gerontology Series B: Psychological Sciences and Social Sciences, *59*(3), P130–P134. https://doi.org/10.1093/geronb/59.3.P130

Luders, E., Gaser, C., Narr, K. L., & Toga, A. W. (2009). Why Sex Matters: Brain Size Independent Differences in Gray Matter Distributions between Men and Women. *The Journal of Neuroscience*, *29*(45), 14265–14270. https://doi.org/10.1523/JNEUROSCI.2261-09.2009

Luders, E., Toga, A. W., & Thompson, P. M. (2014). Why size matters: Differences in brain volume account for apparent sex differences in callosal anatomy: The sexual dimorphism of the corpus callosum. *NeuroImage*, *84*, 820–824. https://doi.org/10.1016/j.neuroimage.2013.09.040

Lundervold, A. J., Wollschläger, D., & Wehling, E. (2014). Age and sex related changes in episodic memory function in middle aged and older adults. *Scandinavian Journal of Psychology*, *55*(3), 225–232. https://doi.org/10.1111/sjop.12114

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford University Press.

Maeda, Y., & Yoon, S. Y. (2013). A Meta-Analysis on Gender Differences in Mental Rotation Ability Measured by the Purdue Spatial Visualization Tests: Visualization of Rotations (PSVT:R). *Educational Psychology Review*, *25*(1), 69–94. https://doi.org/10.1007/s10648-012-9215-x

Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., Poewe, W., Mollenhauer, B., Klinik, P.-E., Sherer, T., Frasier, M., Meunier, C., Rudolph, A., Casaceli, C., Seibyl, J., … Taylor, P. (2011a). The Parkinson Progression Marker Initiative (PPMI).

Progress in Neurobiology, *95*(4), 629–635. https://doi.org/10.1016/j.pneurobio.2011.09.005

Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., Flagg, E., Chowdhury, S., Poewe, W., Mollenhauer, B., Klinik, P.-E., Sherer, T., Frasier, M., Meunier, C., Rudolph, A., Casaceli, C., Seibyl, J., … Taylor, P. (2011b). The Parkinson Progression Marker Initiative (PPMI). *Progress in Neurobiology*, *95*(4), 629–635. https://doi.org/10.1016/j.pneurobio.2011.09.005

Mayo, C. D., Mazerolle, E. L., Ritchie, L., Fisk, J. D., Gawryluk, J. R., & Alzheimer's Disease Neuroimaging Initiative. (2017). Longitudinal changes in microstructural white matter metrics in Alzheimer's disease. *NeuroImage. Clinical*, *13*, 330–338. https://doi.org/10.1016/j.nicl.2016.12.012

Mazure, C. M., & Jones, D. P. (2015). Twenty years and still counting: Including women as participants and studying sex and gender in biomedical research. *BMC Women's Health*, *15*(1). https://doi.org/10.1186/s12905-015-0251-9

McCarrey, A. C., An, Y., Kitner-Triolo, M. H., Ferrucci, L., & Resnick, S. M. (2016). Sex differences in cognitive trajectories in clinically normal older adults. *Psychology and Aging*, *31*(2), 166–175. https://doi.org/10.1037/pag0000070

McDowell, I., Xi, G., Lindsay, J., & Tuokko, H. (2004). Canadian Study of Health and Aging: Study Description and Patterns of Early Cognitive Decline. *Aging, Neuropsychology, and Cognition*, *11*(2–3), 149–168. https://doi.org/10.1080/13825580490511044

McHugh, R. K., Votaw, V. R., Sugarman, D. E., & Greenfield, S. F. (2018). Sex and gender differences in substance use disorders. *Clinical Psychology Review*, *66*, 12–23. https://doi.org/10.1016/j.cpr.2017.10.012

McLean, C. P., Asnaani, A., Litz, B. T., & Hofmann, S. G. (2011). Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. *Journal of Psychiatric Research*, *45*(8), 1027–1035. https://doi.org/10.1016/j.jpsychires.2011.03.006

Menzler, K., Belke, M., Wehrmann, E., Krakow, K., Lengler, U., Jansen, A., Hamer, H. M., Oertel, W. H., Rosenow, F., & Knake, S. (2011). Men and women are different: Diffusion tensor imaging reveals sexual dimorphism in the microstructure of the thalamus, corpus callosum and cingulum. *NeuroImage*, *54*(4), 2557–2562. https://doi.org/10.1016/j.neuroimage.2010.11.029

Miller, J. L., Weiss, L. G., Beal, A. L., Saklofske, D. H., Zhu, J., & Holdnack, J. A. (2015). Intelligent Use of Intelligence Tests: Empirical and Clinical Support for Canadian WAIS–IV Norms. *Journal of Psychoeducational Assessment*, *33*(4), 312–328. https://doi.org/10.1177/0734282915578577

Monsalves, M. J., Bangdiwala, A. S., Thabane, A., & Bangdiwala, S. I. (2020). LEVEL (Logical Explanations & Visualizations of Estimates in Linear mixed models): Recommendations for reporting multilevel data and analyses. *BMC Medical Research Methodology*, *20*(1), 3. https://doi.org/10.1186/s12874-019-0876-8

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the*

American Geriatrics Society, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

National Institutes of Health. (n.d.). *Sex & Gender | Office of Research on Women's Health*. Retrieved June 9, 2022, from https://orwh.od.nih.gov/sex-gender

Ngun, T. C., Ghahramani, N., Sánchez, F. J., Bocklandt, S., & Vilain, E. (2011). The genetics of sex differences in brain and behavior. *Frontiers in Neuroendocrinology*, *32*(2), 227–246. https://doi.org/10.1016/j.yfrne.2010.10.001

Ohlhauser, L., Parker, A. F., Smart, C. M., & Gawryluk, J. R. (2018). White matter and its relationship with cognition in subjective cognitive decline. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*, *11*, 28–35. https://doi.org/10.1016/j.dadm.2018.10.008

Ouyang, Y., Cui, D., Yuan, Z., Liu, Z., Jiao, Q., Yin, T., & Qiu, J. (2021). Analysis of Age-Related White Matter Microstructures Based on Diffusion Tensor Imaging. *Frontiers in Aging Neuroscience*, *13*, 664911. https://doi.org/10.3389/fnagi.2021.664911

Park, D. C., & Reuter-Lorenz, P. (2009). The Adaptive Brain: Aging and Neurocognitive Scaffolding. *Annual Review of Psychology*, *60*(1), 173–196. https://doi.org/10.1146/annurev.psych.59.103006.093656

Parker, A. F., Ohlhauser, L., Scarapicchia, V., Smart, C. M., Szoeke, C., & Gawryluk, J. R. (2022). A Systematic Review of Neuroimaging Studies Comparing Individuals with Subjective Cognitive Decline to Healthy Controls. *Journal of Alzheimer's Disease*, *86*(4), 1545–1567. https://doi.org/10.3233/JAD-215249

Plevkova, J., Brozmanova, M., Harsanyiova, J., Sterusky, M., Honetschlager, J., & Buday, T. (2021). Various Aspects of Sex and Gender Bias in Biomedical Research. *Physiological Research*, S367–S378. https://doi.org/10.33549/physiolres.934593

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *The American Journal of Psychiatry*, *164*(6), 942–948. https://doi.org/10.1176/ajp.2007.164.6.942

Potvin, O., Mouiha, A., Dieumegarde, L., & Duchesne, S. (2018). Corrigendum to 'Normative data for subcortical regional volumes over the lifetime of the adult human brain' [NeuroImage 137 (2016) 9–20]. *NeuroImage*, *183*, 994–995. https://doi.org/10.1016/j.neuroimage.2018.09.020

Potvin, O., Mouiha, A., Dieumegarde, L., Duchesne, S., & Alzheimer's Disease Neuroimaging Initiative. (2016). Normative data for subcortical regional volumes over the lifetime of the adult human brain. *NeuroImage*, *137*, 9–20. https://doi.org/10.1016/j.neuroimage.2016.05.016

Proust-Lima, C., Amieva, H., Letenneur, L., Orgogozo, J.-M., Jacqmin-Gadda, H., & Dartigues, J.-F. (2008). Gender and education impact on brain aging: A general cognitive factor approach. *Psychology and Aging*, *23*(3), 608–620. https://doi.org/10.1037/a0012838

Raman, R., Quiroz, Y. T., Langford, O., Choi, J., Ritchie, M., Baumgartner, M., Rentz, D., Aggarwal, N. T., Aisen, P., Sperling, R., & Grill, J. D. (2021). Disparities by Race and Ethnicity Among Adults Recruited for a Preclinical Alzheimer Disease

Trial. *JAMA Network Open*, *4*(7), e2114364. https://doi.org/10.1001/jamanetworkopen.2021.14364

Raz, N., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Lindenberger, U. (2010). Trajectories of brain aging in middle-aged and older adults: Regional and individual differences. *NeuroImage*, *51*(2), 501–511. https://doi.org/10.1016/j.neuroimage.2010.03.020

Raz, N., & Rodrigue, K. M. (2006). Differential aging of the brain: Patterns, cognitive correlates and modifiers. *Neuroscience & Biobehavioral Reviews*, *30*(6), 730–748. https://doi.org/10.1016/j.neubiorev.2006.07.001

Raz, N., Rodrigue, K. M., Kennedy, K. M., & Acker, J. D. (2007). Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. *Neuropsychology*, *21*(2), 149–157. https://doi.org/10.1037/0894-4105.21.2.149

Reeve, A., Simcox, E., & Turnbull, D. (2014). Ageing and Parkinson's disease: Why is advancing age the biggest risk factor? *Ageing Research Reviews*, *14*, 19–30. https://doi.org/10.1016/j.arr.2014.01.004

Reitz, C., & Mayeux, R. (2014). Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers. *Biochemical Pharmacology*, *88*(4), 640–651. https://doi.org/10.1016/j.bcp.2013.12.024

Reuter-Lorenz, P. A., & Park, D. C. (2014). How Does it STAC Up? Revisiting the Scaffolding Theory of Aging and Cognition. *Neuropsychology Review*, *24*(3), 355–370. https://doi.org/10.1007/s11065-014-9270-9

Rich-Edwards, J. W., Kaiser, U. B., Chen, G. L., Manson, J. E., & Goldstein, J. M. (2018). Sex and Gender Differences Research Design for Basic, Clinical, and

Population Studies: Essentials for Investigators. *Endocrine Reviews*, *39*(4), 424–439. https://doi.org/10.1210/er.2017-00246

Rippon, G., Jordan-Young, R., Kaiser, A., & Fine, C. (2014). Recommendations for sex/gender neuroimaging research: Key principles and implications for research design, analysis and interpretation. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00650

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., Harris, M. A., Alderson, H. L., Hunter, S., Neilson, E., Liewald, D. C. M., Auyeung, B., Whalley, H. C., Lawrie, S. M., Gale, C. R., Bastin, M. E., McIntosh, A. M., & Deary, I. J. (2018). Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. *Cerebral Cortex*, *28*(8), 2959–2975. https://doi.org/10.1093/cercor/bhy109

Rorden, C., & Brett, M. (2000). Stereotaxic Display of Brain Lesions. *Behavioural Neurology*, *12*(4), 191–200. https://doi.org/10.1155/2000/421719

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology*, *77*(13), 1272–1275. https://doi.org/10.1212/WNL.0b013e318230208a

Roy, D. (2012). Neuroethics, Gender and the Response to Difference. *Neuroethics*, *5*(3), 217–230. https://doi.org/10.1007/s12152-011-9130-8

Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., & Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Transactions on Medical Imaging*, *18*(8), 712–721. https://doi.org/10.1109/42.796284

Ruigrok, A. N. V., Salimi-Khorshidi, G., Lai, M.-C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience & Biobehavioral Reviews*, *39*, 34–50. https://doi.org/10.1016/j.neubiorev.2013.12.004

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H., & Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *JAMA*, *291*(16), 2007–2012. https://doi.org/10.1001/jama.291.16.2007

Ryan, J., Woods, R. L., Murray, A. M., Shah, R. C., Britt, C. J., Reid, C. M., Wolfe, R., Nelson, M. R., Lockery, J. E., Orchard, S. G., Trevaks, R. E., Chong, T. J., McNeil, J. J., Storey, E., & on behalf of the ASPREE Investigator Group. (2021). Normative performance of older individuals on the Hopkins Verbal Learning Test-Revised (HVLT-R) according to ethno-racial group, gender, age and education level. *The Clinical Neuropsychologist*, *35*(6), 1174–1190. https://doi.org/10.1080/13854046.2020.1730444

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, *143*(8), 783–822. https://doi.org/10.1037/bul0000102

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society : JINS*, *16*(5), 754–760. https://doi.org/10.1017/S1355617710000706

Salthouse, T. A. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, *34*, 17–24. https://doi.org/10.1037/pag0000288

Salthouse, T. A., Schroeder, D. H., & Ferrer, E. (2004). Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Developmental Psychology*, *40*(5), 813–822. https://doi.org/10.1037/0012-1649.40.5.813

Sánchez-Izquierdo, M., & Fernández-Ballesteros, R. (2021). Cognition in Healthy Aging. *International Journal of Environmental Research and Public Health*, *18*(3), 962. https://doi.org/10.3390/ijerph18030962

Sanchis-Segura, C., Ibañez-Gual, M. V., Adrián-Ventura, J., Aguirre, N., Gómez-Cruz, Á. J., Avila, C., & Forn, C. (2019). Sex differences in gray matter volume: How many and how large are they really? *Biology of Sex Differences*, *10*(1), N.PAG-N.PAG. https://doi.org/10.1186/s13293-019-0245-7

Seeman, T. E., Lusignolo, T. M., Albert, M., & Berkman, L. (2001). Social relationships, social support, and patterns of cognitive aging in healthy, high-functioning older adults: MacArthur Studies of Successful Aging. *Health Psychology*, *20*(4), 243–255. https://doi.org/10.1037/0278-6133.20.4.243

Sexton, C. E., Walhovd, K. B., Storsve, A. B., Tamnes, C. K., Westlye, L. T., Johansen-Berg, H., & Fjell, A. M. (2014). Accelerated Changes in White Matter Microstructure during Aging: A Longitudinal Diffusion Tensor Imaging Study. *The Journal of Neuroscience*, *34*(46), 15425–15436. https://doi.org/10.1523/JNEUROSCI.0203-14.2014

Shaked, D., Leibel, D. K., Katzel, L. I., Davatzikos, C., Gullapalli, R. P., Seliger, S. L., Erus, G., Evans, M. K., Zonderman, A. B., & Waldstein, S. R. (2019). Disparities in Diffuse Cortical White Matter Integrity Between Socioeconomic Groups.

Frontiers in Human Neuroscience, *13*, 198–198. https://doi.org/10.3389/fnhum.2019.00198

Shiino, A., Chen, Y., Tanigaki, K., Yamada, A., Vigers, P., Watanabe, T., Tooyama, I., & Akiguchi, I. (2017). Sex-related difference in human white matter volumes studied: Inspection of the corpus callosum and other white matter by VBM. *Scientific Reports*, *7*(1), 39818. https://doi.org/10.1038/srep39818

Singer, T., Verhaeghen, P., Ghisletta, P., Lindenberger, U., & Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology and Aging*, *18*(2), 318–331. https://doi.org/10.1037/0882-7974.18.2.318

Smith. (1973). *Symbol Digit Modalities Test*. Western Psychological Services.

Smith. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155. https://doi.org/10.1002/hbm.10062

Smith. (2005). Relative Size versus Controlling for Size: Interpretation of Ratios in Research on Sexual Dimorphism in the Human Corpus Callosum. *Current Anthropology*, *46*(2), 249–273. https://doi.org/10.1086/427117

Smith, Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004a). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl 1*, S208-219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Smith, Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004b). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., & Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*(4), 1487–1505. https://doi.org/10.1016/j.neuroimage.2006.02.024

Smith, S., & Nichols, T. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, *44*(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Soares, J. M., Marques, P., Alves, V., & Sousa, N. (2013). A hitchhiker’s guide to diffusion tensor imaging. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00031

Strauss, E., Sherman, E. M. S., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed). Oxford University Press.

Strober, L. B., Bruce, J. M., Arnett, P. A., Alschuler, K. N., Lebkuecher, A., Di Benedetto, M., Cozart, J., Thelen, J., Guty, E., & Roman, C. (2020). A new look at an old test: Normative data of the symbol digit modalities test –Oral version. *Multiple Sclerosis and Related Disorders*, *43*, 102154. https://doi.org/10.1016/j.msard.2020.102154

Sullivan, E. V., Adalsteinsson, E., Hedehus, M., Ju, C., Moseley, M., Lim, K. O., & Pfefferbaum, A. (2001). Equivalent disruption of regional white matter microstructure in ageing healthy men and women. *Neuroreport*, *12*(1), 99–104.

Szoeke, C., Downie, S. J., Parker, A. F., & Phillips, S. (2021). Sex hormones, vascular factors and cognition. *Frontiers in Neuroendocrinology*, *62*, 100927. https://doi.org/10.1016/j.yfrne.2021.100927

Takao, H., Hayashi, N., & Ohtomo, K. (2014). Sex dimorphism in the white matter: Fractional anisotropy and brain size. *Journal of Magnetic Resonance Imaging*, *39*(4), 917–923. https://doi.org/10.1002/jmri.24225

Tierney, M. C., Curtis, A. F., Chertkow, H., & Rylett, R. J. (2017). Integrating sex and gender into neurodegeneration research: A six-component strategy. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *3*(4), 660–667. https://doi.org/10.1016/j.trci.2017.10.006

United Nations, Department of Economic and Social Affairs, & Population Division. (2019). *World population prospects Highlights, 2019 revision Highlights, 2019 revision*.

United Nations, Department of Economic and Social Affairs, & Population Division. (2020). *World population ageing, 2019 highlights.*

van Hemmen, J., Saris, I. M. J., Cohen-Kettenis, P. T., Veltman, D. J., Pouwels, P. J. W., & Bakker, J. (2017). Sex Differences in White Matter Microstructure in the

Human Brain Predominantly Reflect Differences in Sex Hormone Exposure. *Cerebral Cortex*, *27*(5), 2994–3001. https://doi.org/10.1093/cercor/bhw156

Voevodskaya, O., Simmons, A., Nordenskjöld, R., Kullberg, J., Ahlström, H., Lind, L., Wahlund, L.-O., Larsson, E.-M., & Westman, E. (2014). The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease. *Frontiers in Aging Neuroscience*, *6*. https://doi.org/10.3389/fnagi.2014.00264

Vos, S. B., Jones, D. K., Jeurissen, B., Viergever, M. A., & Leemans, A. (2012). The influence of complex white matter architecture on the mean diffusivity in diffusion tensor MRI of the human brain. *NeuroImage*, *59*(3), 2208–2216. https://doi.org/10.1016/j.neuroimage.2011.09.086

Vos, S. B., Jones, D. K., Viergever, M. A., & Leemans, A. (2011). Partial volume effect as a hidden covariate in DTI analyses. *NeuroImage*, *55*(4), 1566–1576. https://doi.org/10.1016/j.neuroimage.2011.01.048

Voyer, D., Saint Aubin, J., Altman, K., & Gallant, G. (2021). Sex differences in verbal working memory: A systematic review and meta-analysis. *Psychological Bulletin*, *147*(4), 352–398. https://doi.org/10.1037/bul0000320

Voyer, D., Voyer, S. D., & Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. *Psychonomic Bulletin & Review*, *24*(2), 307–334. https://doi.org/10.3758/s13423-016-1085-7

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language*, *108*(3), 175–183. https://doi.org/10.1016/j.bandl.2008.07.001

Wang, H., Lo, M.-T., Rosenthal, S. B., Makowski, C., Andreassen, O. A., Salem, R. M., McEvoy, L. K., McEvoy, L. K., Fiecas, M., & Chen, C.-H. (2021). Similar Genetic Architecture of Alzheimer's Disease and Differential APOE Effect Between Sexes. *Frontiers in Aging Neuroscience*, *13*. https://doi.org/10.3389/fnagi.2021.674318

Wang, L., Shen, H., Tang, F., Zang, Y., & Hu, D. (2012). Combined structural and resting-state functional MRI analysis of sexual dimorphism in the young adult human brain: An MVPA approach. *NeuroImage*, *61*(4), 931–940. https://doi.org/10.1016/j.neuroimage.2012.03.080

Webb, K. L., Ryan, J., Wolfe, R., Woods, R. L., Shah, R. C., Murray, A. M., Orchard, S. G., & Storey, E. (2022). Test-Retest Reliability and Minimal Detectable Change of Four Cognitive Tests in Community-Dwelling Older Adults. *Journal of Alzheimer's Disease*, 1–11. https://doi.org/10.3233/JAD-215564

Weber, D., Skirbekk, V., Freund, I., & Herlitz, A. (2014). The changing face of cognitive gender differences in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(32), 11673–11678. https://doi.org/10.1073/pnas.1319538111

Westerhausen, R., Kompus, K., Dramsdahl, M., Falkenberg, L. E., Grüner, R., Hjelmervik, H., Specht, K., Plessen, K., & Hugdahl, K. (2011). A critical re-examination of sexual dimorphism in the corpus callosum microstructure. *NeuroImage*, *56*(3), 874–880. https://doi.org/10.1016/j.neuroimage.2011.03.013

Wetherell, J. L., Reynolds, C. A., Gatz, M., & Pedersen, N. L. (2002). Anxiety, Cognitive Performance, and Cognitive Decline in Normal Aging. *The Journals of*

Gerontology Series B: Psychological Sciences and Social Sciences, *57*(3), P246–P255. https://doi.org/10.1093/geronb/57.3.P246

Williams, C. M., Peyre, H., Toro, R., & Ramus, F. (2021). Sex differences in the brain are not reduced to differences in body size. *Neuroscience & Biobehavioral Reviews*, *130*, 509–511. https://doi.org/10.1016/j.neubiorev.2021.09.015

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E. (2014). Permutation inference for the general linear model. *NeuroImage*, *92*, 381–397. https://doi.org/10.1016/j.neuroimage.2014.01.060

Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., & Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, *45*(1 Suppl), S173-186. https://doi.org/10.1016/j.neuroimage.2008.10.055

World Health Organization. (n.d.). *Gender and health*. Retrieved June 9, 2022, from http://www.who.int/health-topics/gender

World Health Organization. (2020). *World health statistics 2020: Monitoring health for the SDGs, sustainable development goals*. World Health Organization. https://apps.who.int/iris/handle/10665/332070

Yap, Q. J., Teh, I., Fusar-Poli, P., Sum, M. Y., Kuswanto, C., & Sim, K. (2013). Tracking cerebral white matter changes across the lifespan: Insights from diffusion tensor imaging studies. *Journal of Neural Transmission*, *120*(9), 1369–1395. https://doi.org/10.1007/s00702-013-0971-7

Zagni, E., Simoni, L., & Colombo, D. (2016). Sex and Gender Differences in Central Nervous System-Related Disorders. *Neuroscience Journal*, *2016*. https://doi.org/10.1155/2016/2827090

Zeng, Y., Nie, C., Min, J., Chen, H., Liu, X., Ye, R., Chen, Z., Bai, C., Xie, E., Yin, Z., Lv, Y., Lu, J., Li, J., Ni, T., Bolund, L., Land, K. C., Yashin, A., O'Rand, A. M., Sun, L., … Vaupel, J. (2018). Sex Differences in Genetic Associations With Longevity. *JAMA Network Open*, *1*(4), e181670. https://doi.org/10.1001/jamanetworkopen.2018.1670

Zhang, Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424

Zhang, J., Zhou, W., Wang, L., & Zhang, X. (2017). Gender differences of neuropsychological profiles in cognitively normal older people without amyloid pathology. *Comprehensive Psychiatry*, *75*, 22–26. https://doi.org/10.1016/j.comppsych.2017.02.008

Appendix A. Initial Baseline Inclusion/Exclusion Criteria for Healthy Control Participants from the Parkinson Progression Marker's Initiative

Retrieved from the PPMI Database (STUDY DOCS → Study Protocol & CRFs → "Study CRFs & Assessments"

Healthy Control Inclusion Criteria:

1. Ability to provide written informed consent in accordance with Good Clinical Practice (GCP), International Conference on Harmonization (ICH), and local regulations.
2. Willing and able to comply with scheduled visits, required study procedures and laboratory tests.
3. Women may not be pregnant, lactating, or planning pregnancy during the course of the study.
4. Male or female age 30 years or older at screening.

Healthy Control Exclusion Criteria:

1. Received any of the following drugs that might interfere with dopamine transporter SPECT imaging: Neuroleptics, metoclopramide, alpha methyldopa, methylphenidate, reserpine, or amphetamine derivative, within 6 months of Screening.
2. Current treatment with anticoagulants (e.g., coumadin, heparin) that might preclude safe completion of the lumbar puncture.
3. Condition that precludes the safe performance of routine lumbar puncture, such as prohibitive lumbar spinal disease, bleeding diathesis, or clinically significant coagulopathy or thrombocytopenia.
4. Any other medical or psychiatric condition or lab abnormality, which in the opinion of the investigator might preclude participation.
5. Use of investigational drugs or devices within 60 days prior to Baseline (dietary supplements taken outside of a clinical trial are not exclusionary, e.g., coenzyme Q10).
6. Previously obtained MRI scan with evidence of clinically significant neurological disorder (in the opinion of the Investigator).
7. Current or active clinically significant neurological disorder (in the opinion of the Investigator).
8. First degree relative with idiopathic PD (parent, sibling, child).
9. MoCA score less than or equal to 26.

Note. The above shows the initial inclusion and exclusion at the baseline time point. All following time points used the following exclusion criteria:

1. Two or more neuropsychological test scores < 1.5 standard deviations below the mean.
2. A PPMI investigator rating of mild cognitive impairment or dementia.

Appendix B. Study 3 Syntax for Model Specification in R using lme4

Coding Note:

ID= Identification variable for each participant, time is nested within individuals

Level-1 predictor:

time = continuous variable of years since baseline assessment

Level-2 predictors:

sex = categorical variable, 0=female, 1=male
age = continuous variable of age in years, centered on mean age at baseline
edu = continuous variable of education in years, centered on mean education
NP = neuropsychological outcome measure

Seven final models were created using each of the following neuropsychological outcome measures:

1. Montreal Cognitive Assessment
2. Letter Number Sequencing
3. Semantic Fluency Test
4. Symbol Digit Modalities Test
5. Benton Judgment of Line Orientation Test
6. Hopkins Verbal Learning Test - Revised - Immediate Recall
7. Hopkins Verbal Learning Test - Revised - Delayed Recall

Six models were summarized in Appendix C - Tables 4.5-4.11 for each of the seven neuropsychological measures:

Model 1 (intercept only) = NP_m1
Model 2 (time only) = MP_m2
Model 3 (age, sex, and education as level-2 predictors) = NP_m5_age_edu_sex
Model 4 (time X sex interaction) = NP_m6_timeXsex
Model 5 (time X age interaction) = NP_m7_timeXage
Model 6 (both time X sex and time X age interactions) = NP_timeXage_timeXsex

A summary of all final models is found in Table 4.4.

Syntax for Model Specification in R using lme4:

```
#load packages
library(lme4)
library(sjPlot)

##NP models###
#UNCONDITIONAL MODEL (i.e., intercept-only model with no predictors)
NP_m1<-lmer(NP~ 1 + (1 | ID), data=data)
summary(NP_m1)
tab_model(NP_m1)

#UNCONDITIONAL GROWTH MODEL
#adding in time as level-1 predictor
NP_m2<-lmer(NP ~ time + (time | ID), data=data)
summary(NP_m2)
tab_model(NP_m1,NP_m2)
anova(NP_m1,NP_m2)

##adding in baseline age, centered on the mean age at BL
NP_m3_age<-lmer(NP ~ time + age + (time | ID), data=data)
summary(NP_m3_age)
tab_model(NP_m1,NP_m2,NP_m3_age)
anova(NP_m1,NP_m2,NP_m3_age)

## adding in baseline education, centered on the mean education at BL
NP_m4_age_edu<-lmer(NP ~ time + age + edu + (time | ID), data=data)
summary(NP_m4_age_edu)
tab_model(NP_m1,NP_m2,NP_m3,NP_m4_age_edu)
anova(NP_m3_age,NP_m4_age_edu)

##adding in sex, 0=female, 1=male
NP_m5_age_edu_sex<-lmer(NP ~ time + sex + age + edu + (time | ID), data=data)
summary(NP_m5_age_edu_sex)
tab_model(NP_m3_age,NP_m4_age_edu,NP_m5_age_edu_sex)
anova(NP_m3_age,NP_m4_age_edu,NP_m5_age_edu_sex)

#CONDITIONAL GROWTH MODEL = interaction between time and sex
NP_m6_timeXsex<-lmer(NP ~ time * sex + age + edu + (time | ID), data=data)
summary(NP_m6_timeXsex)
tab_model(NP_m1,NP_m5_age_edu_sex,NP_m6_timeXsex)
anova(NP_m1,NP_m5_age_edu_sex,NP_m6_timeXsex)

#CONDITIONAL GROWTH MODEL = interaction between time and BL age
NP_m7_timeXage<-lmer(NP ~time*age + edu + (time | ID), data=data)
```

```
summary(NP_m7_timeXage)
tab_model(NP_m1,NP_m5_age_edu_sex,NP_m7_timeXage)
anova(NP_m1,NP_m5_age_edu_sex,NP_m7_timeXage)

#CONDITIONAL GROWTH MODEL --> 2 interactions: sex*time and edu*time
NP_timeXage_timeXsex<-lmer(NP ~time*age + time*sex+edu + (time | ID), data=data)
summary(NP_timeXage_timeXsex)
tab_model(NP_m1,NP_m5_age_edu_sex,NP_m6_timeXsex,NP_m7_timeXage,NP_timeXage_timeXsex)
anova(NP_m1,NP_m5_age_edu_sex,NP_m6_timeXsex,NP_m7_timeXage,NP_timeXage_timeXsex)
```

Appendix C. Study 3 Description of Final Model Fit and Specification for Each Neuropsychological Outcome Measure

Montreal Cognitive Assessment

Estimates and model fit indices for each iteration of the model to predict MoCA scores can be found in Table 4.5. The intercept only model (Model 1 in Table 4.5) showed significant variability in MoCA scores, and the ICC was 0.47. Adding time as a level-1 predictor (Model 2) was significant and the ICC increased to 0.53. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3). The addition of time by sex interaction (Model 4) did not significantly improve the fit of the model. However, the addition of a time by age interaction (Model 5) did improve the model fit. The final model (Model 6) with both interactions (time by sex and time by age) did not significantly differ in values of the model fit indices and was thus used as the final model. This demonstrated that MoCA scores systematically differed by baseline age. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, MoCA scores declined by 0.03 points. There was a very small interaction of change in MoCA scores over time and baseline age, where higher baseline age was predictive of steeper declines in MoCA scores. The ICC in the final model indicated significant clustering of scores within-person, where neuropsychological scores within an individual were highly related over time (i.e., ICC=0.74 in the final model). In other words, the ICC reflects significant between-person variance in score.

Table 4.5

Results of Multilevel Models to Predict Montreal Cognitive Assessment Scores

Predictors	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5***	**Model 6**
(Intercept)	27.56 (27.34 – 27.7) **p<0.001**	27.75 (27.55 – 27.95) **p<0.001**	27.94 (27.62 – 28.26) **p<0.001**	27.91 (27.58 – 28.23) p**<0.001**	27.76 (27.56 – 27.95) p**<0.001**	27.92 (27.59 – 28.24) p**<0.001**
Time		-0.09 (-0.16 – -0.03) p=**0.005**	-0.09 (-0.15 – -0.03) p=**0.006**	-0.03 (-0.14 – 0.07) p=0.511	-0.09 (-0.15 – -0.03) p=**0.005**	-0.04 (-0.15 – 0.06) p=0.394
Sex			-0.3 (-0.69 – 0.10) p=0.146	-0.23 (-0.64 – 0.17) p=0.261		-0.24 (-0.65 – 0.16) p=0.241
Age			-0.04 (-0.06 – -0.02) p**<0.001**	-0.04 (-0.06 – -0.02) p**<0.001**	-0.03 (-0.05 – -0.02) p**<0.001**	-0.03 (-0.05 – -0.02) p**<0.001**
Education			0.01 (0.06 – 0.07) p=0.819	0.01 (-0.06 – 0.07) p=0.816	0 (-0.07 – 0.07) p=0.988	0.01 (-0.06 – 0.07) p=0.824
Time X Sex				-0.09 (-0.22 – 0.04) p=0.179		-0.08 (-0.21 – 0.06) p=0.261
Time X Age					-0.01 (-0.01 – -0.00) p=**0.005**	-0.01 (-0.01 – -0.00) p=**0.007**
Random Effects						
σ^2	2.31	2.08	2.1	2.1	2.08	2.08
τ_{00}	2.05 $_{\text{ID}}$	0.85 $_{\text{ID}}$	0.70 $_{\text{ID}}$	0.70 $_{\text{ID}}$	0.73 $_{\text{ID}}$	0.72 $_{\text{ID}}$
τ_{11}		0.06 $_{\text{ID.Time}}$	0.05 $_{\text{ID.Time}}$	0.05 $_{\text{ID.Time}}$	0.05 $_{\text{ID.Time}}$	0.05 $_{\text{ID.Time}}$
ρ_{01}		1.00 $_{\text{ID}}$	1.00 $_{\text{ID}}$	1.00 $_{\text{ID}}$	1.00 $_{\text{ID}}$	1.00 $_{\text{ID}}$
ICC	0.47	0.53	0.75	0.75	0.74	0.74
Marginal R^2 / Conditional R^2	0.000 / 0.471	0.006 / 0.537	0.098 / NA	0.113 / NA	0.158 / NA	0.172 / NA
AIC	4002.7	3937.6	3922.3	3922.5	3916.6	3917.2
BIC	4017.4	3967.1	3966.5	3971.6	3960.8	3971.2

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Letter Number Sequencing

Estimates and model fit indices for each iteration of the model to predict LNS scores can be found in Table 4.6. The intercept only model (Model 1 in Table 4.6), showed significant variability in LNS scores and the ICC was 0.66. Adding time as a level-1 predictor (Model 2) was significant and the ICC increased to 0.53. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with further reductions in AIC and BIC (Model 3). The addition of time by sex interaction (Model 4) and a time by age interaction (Model 5) did not significantly improve the model. The final model (Model 6) with both interactions (time by sex and time by age) did not significantly differ in values of the model fit indices and Model 6 was used as the final model. This demonstrated that LNS did not systematically change over time. Baseline scores on the LNS did not significantly differ by sex at the baseline assessment. Higher age at baseline was predictive of lower scores on LNS at baseline. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, LNS scores declined by 0.08 points. There was no interaction between change in LNS scores over time and baseline age or between change in LNS scores over time and sex.

Table 4.6

Results of Multilevel Models to Predict Letter Number Sequencing Scores

Predictors	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5***	**Model 6**
(Intercept)	10.88 (10.56 – 11.20) p**<0.001**	10.91 (10.58 – 11.24) **p <0.001**	10.98 (10.48 – 11.48) p**<0.001**	10.87 (10.36 – 11.39) p**<0.001**	10.93 (10.62 – 11.23) p**<0.001**	10.88 (10.36 – 11.39) p**<0.001**
Time		-0.02 (-0.08 – 0.05) p=0.636	-0.01 (0.08 – 0.05) p=0.675	0.06 (-0.05 – 0.16) p=0.293	-0.01 (0.08 – 0.05) p=0.669	0.05 (-0.05 – 0.16) p=0.324
Sex			-0.08 (-0.69 – 0.53) p=0.795	0.09 (-0.56 – 0.73) p=0.789		0.08 (-0.56 – 0.73) p=0.806
Age			-0.08 (-0.11 – -0.06) p**<0.001**	-0.08 (-0.11 – -0.06) p**<0.001**	-0.08 (-0.10 – -0.05) p**<0.001**	-0.08 (-0.11 – -0.05) p**<0.001**
Education			0.15 (0.05 – 0.25) p=**0.004**	0.15 (0.05 – 0.25) p=**0.005**	0.15 (0.05 – 0.25) p=**0.005**	0.15 (0.05 – 0.25) p=**0.005**
Time X Sex				-0.11 (-0.25 – 0.02) p=0.096		-0.11 (-0.24 – 0.03) p=0.112
Time X Age					0 (-0.01 – 0.00) p=0.34	0 (-0.01 – 0.00) p=0.412
Random Effects						
σ^2	2.41	2.25	2.25	2.25	2.25	2.25
τ_{00}	4.72_{ID}	4.16_{ID}	3.44_{ID}	3.44_{ID}	3.41_{ID}	3.43_{ID}
τ_{11}		$0.05_{\text{ID.Time}}$	$0.05_{\text{ID.Time}}$	$0.05_{\text{ID.Time}}$	$0.05_{\text{ID.Time}}$	$0.05_{\text{ID.Time}}$
ρ_{01}		0.16_{ID}	0.10_{ID}	0.11_{ID}	0.11_{ID}	0.11_{ID}
ICC	0.66	0.69	0.64	0.64	0.64	0.64
Marginal R^2 / Conditional R^2	0.000 / 0.662	0.000 / 0.685	0.125 / 0.687	0.128 / 0.687	0.132 / 0.688	0.133 / 0.689
AIC	4178	4171.2	4139.6	4138.8	4138.7	4140.1
BIC	4192.7	4200.6	4183.7	4187.8	4182.9	4194.1

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Semantic Fluency Test

Estimates and model fit indices for each iteration of the model to predict SFT scores can be found in Table 4.7. The intercept only model (Model 1 in Table 4.7), showed significant variability in SFT scores and the ICC was 0.77. Adding time as a level-1 predictor (Model 2) was non-significant, suggesting no systematic change in SFT scores within individuals over time. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3), where sex, age, and education were significant predictors of SFT scores. The addition of time by sex interaction (Model 4) and a time by age interaction (Model 5) did not significantly improve the model. The final model estimates (Model 6) with both interactions (time by sex and time by age) did not significantly differ from the values of the best fitting model (Model 3) and was thus shown in the final model summary table (Table 4.4). This demonstrated that SFT scores did not systematically change over time. There were significant sex differences in SFT scores at the baseline assessment, where females provided 7.29 more words on average than males. Higher age at baseline was predictive of lower scores on SFT. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, SFT scores declined by 0.23 words. Higher education also predicted higher SFT scores, where each additional year of education beyond the average of 16 years, increased SFT scores by 0.51. There was no interaction between change in SFT scores over time and baseline age or between change in SFT scores over time and sex

Table 4.7

Results of Multilevel Models to Predict Semantic Fluency Scores

Predictors	**Model 1**	**Model 2**	**Model 3***	**Model 4**	**Model 5**	**Model 6**
(Intercept)	52.19 (50.72 – 53.67) p<**0.001**	52.31 (50.81 – 53.81) **p <0.001**	57.06 (54.80 – 59.32) p<**0.001**	57.02 (54.71 – 59.33) p<**0.001**	52.37 (50.91 – 53.82) p<**0.001**	57.04 (54.73 – 59.35) p<**0.001**
Time		-0.05 (-0.29 – 0.20) p=0.699	-0.05 (-0.29 – 0.19) p=0.692	-0.02 (-0.42 – 0.37) p=0.91	-0.05 (-0.29 – 0.20) p=0.711	-0.04 (-0.43 – 0.36) p=0.861
Sex			-7.32 (-10.12 – -4.52) p<**0.001**	-7.26 (-10.16 – -4.36) p<**0.001**		-7.29 (-10.19 – -4.39) p<**0.001**
Age			-0.24 (-0.36 – -0.12) p<**0.001**	-0.24 (-0.36 – -0.12) p<**0.001**	-0.25 (-0.38 – -0.12) p<**0.001**	-0.23 (-0.35 – -0.10) p<**0.001**
Education			0.51 (0.04 – 0.98) p=**0.032**	0.51 (0.04 – 0.98) p=**0.032**	0.35 (-0.15 – 0.85) p=0.166	0.51 (0.04 – 0.98) p=**0.033**
Time X Sex				-0.04 (-0.55 – 0.46) p=0.869		-0.02 (-0.53 – 0.48) p=0.929
Time X Age					-0.01 (-0.03 – 0.01) p=0.359	-0.01 (-0.03 – 0.01) p=0.371
Random Effects						
σ^2	30.58	27.49	27.48	27.48	27.48	27.47
τ_{00}	103.18 $_{ID}$	98.33 $_{ID}$	79.70 $_{ID}$	79.74 $_{ID}$	91.28 $_{ID}$	79.70 $_{ID}$
τ_{11}		0.92 $_{ID.Time}$	0.92 $_{ID.Time}$	0.94 $_{ID.Time}$	0.93 $_{ID.Time}$	0.95 $_{ID.Time}$
ρ_{01}		0.02 $_{ID}$	-0.01 $_{ID}$	-0.02 $_{ID}$	-0.01 $_{ID}$	-0.01 $_{ID}$
ICC	0.77	0.80	0.76	0.76	0.78	0.76
Marginal R^2 / Conditional R^2	0.000 / 0.771	0.000 / 0.796	0.157 / 0.798	0.158 / 0.798	0.071 / 0.798	0.161 / 0.799
AIC	6812.5	6802.4	6766.3	6768.3	6790.7	6769.5
BIC	6827.2	6831.8	6810.5	6817.4	6834.8	6823.5

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Symbol Digit Modalities Test

Estimates and model fit indices for each iteration of the model to predict SDMT scores can be found in Table 4.8. The intercept only model (Model 1 in Table 4.8), showed significant variability in SDMT scores and the ICC was 0.76. Adding time as a level-1 predictor (Model 2) was non-significant, indicating no systematic change in SDMT scores within individuals over time. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3), where sex, age, and education were significant predictors of SDMT scores. The addition of time by sex interaction (Model 4) and a time by age interaction (Model 5) did not significantly improve the model. The final model estimates (Model 6) with both interactions (time by sex and time by age) did not significantly differ from the values of the best fitting model (Model 3) and was thus shown in the final model summary table (Table 4.4). This demonstrated that SDMT scores did not systematically change within individuals over time. There were significant sex differences in SDMT scores at the baseline assessment, where females score 2.83 points higher on average than males. Higher age at baseline was predictive of lower scores on the SDMT. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, SDMT scores declined by 0.5 points. Higher education at baseline was predictive of higher scores on the SDMT. For every additional year of education beyond the mean education years of approximately 16, SDMT scores increased by 0.71 points. There was no significant interaction between change in SDMT scores over time and baseline age or between change in SDMT scores over time and sex.

Table 4.8

Results of Multilevel Models to Predict Symbol Digit Modalities Scores

Predictors	**Model 1**	**Model 2**	**Model 3***	**Model 4**	**Model 5**	**Model 6**
(Intercept)	47.12 (45.74 – 48.49) p<**0.001**	47.05 (45.65 – 48.46) **p <0.001**	48.99 (47.13 – 50.85) p<**0.001**	48.91 (46.98 – 50.84) p<**0.001**	47.17 (46.01 – 48.33) p<**0.001**	48.94 (47.01 – 50.87) p<**0.001**
Time		0.03 (-0.20 – 0.25) p=0.801	0.03 (-0.20 – 0.25) p=0.809	0.07 (-0.29 – 0.43) p=0.709	0.03 (-0.19 – 0.25) p=0.793	0.06 (-0.31 – 0.42) p=0.766
Sex			-2.83 (-5.12 – -0.55) p=**0.015**	-2.72 (-5.14 – -0.30) p=**0.028**		-2.76 (-5.18 – -0.33) p=**0.026**
Age			-0.52 (-0.62 – -0.42) p<**0.001**	-0.52 (-0.62 – -0.42) p<**0.001**	-0.51 (-0.61 – -0.40) p<**0.001**	-0.5 (-0.60 – -0.40) p<**0.001**
Education			0.71 (0.33 – 1.09) p<**0.001**	0.71 (0.33 – 1.09) p<**0.001**	0.65 (0.26 – 1.03) **p=0.001**	0.71 (0.33 – 1.09) p<**0.001**
Time X Sex				-0.07 (-0.53 – 0.39) p=0.776		-0.05 (-0.51 – 0.42) p=0.845
Time X Age					-0.01 (-0.03 – 0.01) p=0.296	-0.01 (-0.03 – 0.01) p=0.302
Random Effects						
σ^2	28.71	26.87	26.88	26.88	26.86	26.86
τ_{00}	89.54 ID	84.59 ID	51.14 ID	51.17 ID	52.57 ID	51.18 ID
τ_{11}		0.56 ID.Time	0.55 ID.Time	0.57 ID.Time	0.56 ID.Time	0.58 ID.Time
ρ_{01}		0.07 ID	-0.03 ID	-0.03 ID	-0.03 ID	-0.04 ID
ICC	0.76	0.77	0.67	0.67	0.68	0.67
Marginal R^2 / Conditional R^2	0.000 / 0.757	0.000 / 0.774	0.307 / 0.773	0.307 / 0.773	0.298 / 0.774	0.312 / 0.775
AIC	6728.8	6726.3	6637.7	6639.7	6642.6	6640.6
BIC	6743.6	6755.8	6681.9	6688.7	6686.8	6694.6

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Benton Judgment of Line Orientation

Estimates and model fit indices for each iteration of the model to predict BJLOT scores can be found in Table 4.9. The intercept only model (Model 1 in Table 4.9), showed significant variability in BJLOT scores and the ICC was 0.65. Adding time as a level-1 predictor (Model 2) was non-significant, indicating no systematic change in BJLOT scores within individuals over time. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3), where sex, age, and education were significant predictors of BJLOT scores. The addition of time by sex interaction (Model 4) was not significant, but the addition of a time by age interaction (Model 5) did significantly improve the model. The final model estimates (Model 6) with both interactions (time by sex and time by age) had slightly lower AIC and slightly higher BIC values compared to the next best fitting model (Model 3) and was thus shown in the final model summary table (Table 4.4). This demonstrated that BJLOT scores did not systematically change within individuals over time. There were significant sex differences in BJLOT scores at the baseline assessment, where males score 1.23 points higher on average than females. Higher age at baseline was predictive of lower scores on the BJLOT. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, BJLOT scores declined by 0.03 points. Higher education at baseline was predictive of higher scores on the BJLOT. For every additional year of education beyond the mean education years of approximately 16, BJLOT scores increased by 0.15 points. There was no significant interaction between change in BJLOT scores over time and sex, but there was a significant interaction between change in BJLOT score and baseline age, where higher age predicted slightly steeper declines in BJLOT scores over time.

Table 4.9

Results of Multilevel Models to Predict Benton Judgement of Line Orientation Scores

Predictors	**Model 1**	**Model 2**	**Model 3***	**Model 4**	**Model 5**	**Model 6**
(Intercept)	12.87 (12.61 – 13.13) p**<0.001**	12.97 (12.70 – 13.24) **p <0.001**	12.23 (11.82 – 12.63) p**<0.001**	12.19 (11.77 – 12.61) p**<0.001**	12.98 (12.72 – 13.24) p**<0.001**	12.2 (11.78 – 12.63) p**<0.001**
Time		-0.05 (-0.10 – 0.01) p=0.103	-0.04 (-0.10 – 0.01) p=0.11	-0.02 (-0.11 – 0.07) p=0.63	-0.05 (-0.10 – 0.01) p=0.09	-0.03 (-0.12 – 0.06) p=0.509
Sex			1.17 (0.68 – 1.67) p**<0.001**	1.23 (0.70 – 1.76) p**<0.001**		1.21 (0.68 – 1.74) p**<0.001**
Age			-0.04 (-0.06 – -0.01) p=**0.001**	-0.04 (-0.06 – -0.01) p=**0.001**	-0.02 (-0.05 – 0.00) p=0.065	-0.03 (-0.05 – -0.00) p=**0.023**
Education			0.15 (0.07 – 0.23) p**<0.001**	0.15 (0.07 – 0.23) p**<0.001**	0.17 (0.09 – 0.26) p**<0.001**	0.15 (0.07 – 0.23) p**<0.001**
Time X Sex				-0.04 (-0.15 – 0.08) p=0.522		-0.03 (-0.14 – 0.09) p=0.658
Time X Age					-0.01 (-0.01 – -0.00) p=**0.012**	-0.01 (-0.01 – -0.00) p=**0.015**
Random Effects						
σ^2	2.13	2.11	2.11	2.11	2.1	2.1
τ_{00}	3.00_{ID}	2.57_{ID}	1.97_{ID}	1.97_{ID}	2.31_{ID}	1.99_{ID}
τ_{11}		$0.00_{\text{ID.Time}}$	$0.00_{\text{ID.Time}}$	$0.00_{\text{ID.Time}}$	$0.00_{\text{ID.Time}}$	$0.00_{\text{ID.Time}}$
ρ_{01}		1.00_{ID}	1.00_{ID}	1.00_{ID}	1.00_{ID}	1.00_{ID}
ICC	0.58	0.55	0.48	0.48	0.56	0.49
Marginal R^2 / Conditional R^2	0.000 / 0.585	0.003 / NA	0.236 / NA	0.232 / NA	0.074 / 0.596	0.247 / NA
AIC	3991.5	3991.0	3956.4	3958	3970.9	3954
BIC	4006.3	4020.4	4000.6	4007.1	4015	4008

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Hopkins Verbal Learning Test – Revised (Immediate Recall)

Estimates and model fit indices for each iteration of the model to predict Immediate Recall scores can be found in Table 4.10. The intercept only model (Model 1 in Table 4.10), showed significant variability in Immediate Recall scores and the ICC was 0.62. Adding time as a level-1 predictor (Model 2) was significant, indicating systematic change in Immediate Recall scores within individuals over time. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3), where sex, age, and education were significant predictors of Immediate Recall scores. The addition of time by sex interaction (Model 4) and a time by age interaction (Model 5) did significantly improve the model. The final model estimates (Model 6) with both interactions (time by sex and time by age) had very similar AIC values and slightly higher BIC values compared to the next best fitting model (Model 3) and was thus shown in the final model summary table (Table 4.4). The final model demonstrated that Immediate Recall scores slightly improved within individuals over time, by approximately 0.25 words per year. There were significant sex differences in Immediate Recall scores at the baseline assessment, such that females recalled 1.23 more words on average compared to males. Higher age at baseline was predictive of lower Immediate Recall scores. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, Immediate Recall scores declined by 0.1 words. Thus a 70-year-old at the baseline assessment would be predicted to recall one fewer word than a 60-year-old at the baseline assessment. Higher education was predictive of higher scores on the Immediate Recall. For every additional year of education beyond the mean education years of approximately 16, Immediate Recall scores increased by 0.28 words. There was no significant interaction between change in Immediate Recall scores over time and sex or between change in Immediate Recall scores and baseline age.

Table 4.10

Results of Multilevel Models to Predict Hopkins Verbal Learning Test - Revised Immediate Recall Scores

Predictors	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6***
(Intercept)	26.26 (25.68 – 26.85) p<**0.001**	25.95 (25.36 – 26.53) p<**0.001**	27.33 (26.44 – 28.23) p<**0.001**	27.15 (26.23 – 28.08) p<**0.001**	25.98 (25.42 – 26.54) p<**0.001**	27.17 (26.24 – 28.09) p<**0.001**
Time		0.13 0.01 – 0.26 p=**0.033**	0.14 (0.01 – 0.26) p=**0.03**	0.26 (0.06 – 0.46) p=**0.01**	0.14 (0.01 – 0.26) p=**0.031**	0.25 (0.05 – 0.45) p=**0.014**
Sex			-2.12 (-3.22 – -1.02) p<**0.001**	-1.84 (-2.99 – -0.68) p=**0.002**		-1.85 (-3.01 – -0.70) p=**0.002**
Age			-0.11 (-0.16 – -0.06) p<**0.001**	-0.11 (-0.16 – -0.06) p<**0.001**	-0.11 (-0.16 – -0.06) p<**0.001**	-0.1 (-0.15 – -0.05) p<**0.001**
Education			0.28 (0.10 – 0.47) p=**0.003**	0.28 (0.10 – 0.47) p=**0.003**	0.23 (0.04 – 0.42) p=**0.015**	0.28 (0.10 – 0.47) p=**0.003**
Time X Sex				-0.2 (-0.45 – 0.05) p=0.121		-0.19 (-0.44 – 0.07) p=0.148
Time X Age					-0.01 (-0.02 – 0.00) p=0.211	-0.01 (-0.02 – 0.00) p=0.254
Random Effects						
σ^2	9.25	8.78	8.78	8.77	8.77	8.76
τ_{00}	15.31 $_{\text{ID}}$	12.26 $_{\text{ID}}$	10.21 $_{\text{ID}}$	10.21 $_{\text{ID}}$	10.86 $_{\text{ID}}$	10.22 $_{\text{ID}}$
τ_{11}		0.12 $_{\text{ID.Time}}$	0.12 $_{\text{ID.Time}}$	0.12 $_{\text{ID.Time}}$	0.13 $_{\text{ID.Time}}$	0.12 $_{\text{ID.Time}}$
ρ_{01}		0.44 $_{\text{ID}}$	0.27 $_{\text{ID}}$	0.28 $_{\text{ID}}$	0.35 $_{\text{ID}}$	0.27 $_{\text{ID}}$
ICC	0.62	0.64	0.59	0.59	0.61	0.59
Marginal R^2 / Conditional R^2	0.000 / 0.623	0.002 / 0.645	0.123 / 0.642	0.131 / 0.645	0.087 / 0.646	0.137 / 0.648
AIC	5491.2	5477.1	5446.2	5445.8	5458.5	5446.5
BIC	5506	5506.5	5490.4	5494.9	5502.7	5500.4

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.07 years. *Indicates best fitting model.

Hopkins Verbal Learning Test – Revised (Delayed Recall)

Estimates and model fit indices for each iteration of the model to predict Delayed Recall scores can be found in Table 11. The intercept only model (Model 1 in Table 10), showed significant variability in Immediate Recall scores and the ICC was 0.61. Adding time as a level-1 predictor (Model 2) was not significant, indicating no systematic change in Delayed Recall scores within individuals over time. The addition of level-2 predictors (age, sex, and education) improved the fit of the model, with reductions in AIC and BIC (Model 3), where sex, age, and education were significant predictors of Immediate Recall scores. The addition of time by sex interaction (Model 4) and a time by age interaction (Model 5) did significantly improve the model. The final model estimates (Model 6) with both interactions (time by sex and time by age) had slightly higher AIC and BIC values compared to the next best fitting model (Model 3) and was thus shown in the final model summary table (Table 4.4). The final model demonstrated that Delayed Recall scores did not systematically change within individuals over time. There were significant sex differences in Delayed Recall scores at the baseline assessment, such that females recalled 0.93 more words on average compared to males. Higher age at baseline was predictive of lower Delayed Recall scores. For every year increase of baseline age beyond the baseline mean age of approximately 60 years, Delayed Recall scores declined by 0.05 words. Higher education was predictive of higher scores on the Delayed Recall. For every additional year of education beyond the mean education years of approximately 16, Delayed Recall scores increased by 0.13 words. There was no significant interaction between change in Immediate Recall scores over time and sex or between change in Immediate Recall scores and baseline age.

Table 4.11

Results of Multilevel Models to Predict Hopkins Verbal Learning Test - Revised Delayed Recall Scores

Predictors	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6***
(Intercept)	9.26 (8.97 – 9.55) p<**0.001**	9.16 (8.86 – 9.46) p<**0.001**	9.79 (9.33 – 10.24) p<**0.001**	9.76 (9.29 – 10.23) p<**0.001**	9.17 (8.89 – 9.46) p<**0.001**	9.77 (9.29 – 10.24) p<**0.001**
Time		0.04 (-0.02 – 0.11) p=0.163	0.04 (-0.02 – 0.10) p=0.174	0.06 (-0.04 – 0.16) p=0.25	0.04 (-0.02 – 0.11) p=0.165	0.05 (-0.05 – 0.15) p=0.288
Sex			-0.96 (-1.52 – -0.40) p=**0.001**	-0.92 (-1.51 – -0.32) p=**0.003**		-0.93 (-1.52 – -0.33) p=**0.002**
Age			-0.06 (-0.08 – -0.04) p<**0.001**	-0.06 (-0.08 – -0.04) p<**0.001**	-0.06 (-0.08 – -0.03) p<**0.001**	-0.05 (-0.08 – -0.03) p<**0.001**
Education			0.13 (0.04 – 0.22) p=**0.007**	0.13 (0.04 – 0.22) p=**0.007**	0.11 (0.01 – 0.20) p=**0.026**	0.13 (0.04 – 0.22) p=**0.007**
Time X Sex				-0.03 (-0.15 – 0.10) p=0.691		-0.02 (-0.15 – 0.11) p=0.764
Time X Age					0 (-0.01 – 0.00) p=0.236	0 (-0.01 – 0.00) p=0.248
Random Effects						
σ^2	2.43	2.34	2.34	2.34	2.34	2.34
τ_{00}	3.76 $_{ID}$	3.29 $_{ID}$	2.68 $_{ID}$	2.68 $_{ID}$	2.85 $_{ID}$	2.68 $_{ID}$
τ_{11}		0.02 $_{ID.Time}$	0.02 $_{ID.Time}$	0.03 $_{ID.Time}$	0.02 $_{ID.Time}$	0.03 $_{ID.Time}$
ρ_{01}		0.27 $_{ID}$	0.23 $_{ID}$	0.21 $_{ID}$	0.25 $_{ID}$	0.22 $_{ID}$
ICC	0.61	0.62	0.57	0.58	0.59	0.58
Marginal R^2 / Conditional R^2	0.000 / 0.607	0.001 / 0.623	0.121 / 0.626	0.122 / 0.627	0.093 / 0.628	0.127 / 0.630
AIC	4143.5	4140.9	4109.3	4111.2	4119.2	4111.8
BIC	4158.2	4170.3	4153.5	4160.2	4163.4	4165.8

Note. ICC=Intraclass Correlation Coefficient. AIC=Akaine Information Criterion (AIC). BIC=Bayesian Information Criterion. Sex was centered so that female=0, male=1. Age was centered on mean baseline age of 60.62 years. Education was centered on mean years of education of 16.01. *Indicates best fitting model.

Appendix D. Data Availability Statement and Funding Details

Data Availability Statement:

Data used in the preparation of this article were obtained from the Parkinson's Progression Markers Initiative (PPMI) database (www.ppmi-info.org/access-dataspecimens/download-data). For up-to-date information on the study, visit ppmi-info.org.

Funding Details:

1. Canadian Institutes of Health Research

This project was supported by the Canada Graduate Scholarship Doctoral Award funded to the author by the Canadian Institutes of Health Research (Competition Number: 201911).

2. Women, Sex, Gender, and Dementia & Canadian Consortium on Neurodegeneration and Aging (CCNA) Team 9

This project was supported by the WSGD & the CCNA Team 9 funded to the author.

3. Funding for the Parkinson's Progression Markers Initiative

This study used publicly available data from the Parkinson's Progression Markers Initiative (PPMI). PPMI – a public-private partnership – is funded by the Michael J. Fox Foundation for Parkinson's Research and funding partners, including: 4D Pharma, AbbVie Inc, AcureX Therapeutic, Allergan, Amathus Therapeutics, Aligning Science Across Parkinson's (ASAP), Avid Radiopharmaceuticals, Bial Biotech, Biogen, BioLegen, Bristol Myers Squibb, Calico Life Sciences LLC, Celgene Corporation, DaCapo Brainscience, Denali Therapeutics, The Edmond J. Safra Foundation, Eli Lilly and Company, GE Healthcare, GlaxoSmithKline, Golub Capital, Handl Therapeutics, Insitro, Janssen Pharmaceuticals, Lundbeck, Merck & Co., Inc., Meso Scale Diagnostics, LLC, Neurocrine Biosciences, Pfizer Inc., Piramal Imaging, Prevail Therapeutics, F. Hoffmann-La Roche Ltd and its affiliated company Genentech Inc., Sanofi Genzyme, Servier, Takeda Pharmaceutical Company, Teva Neuroscience, Inc. , UCB, Vanqua Bio, Verily Life Sciences, Voyager Therapeutics, Inc., Yumanity Therapeutics, Inc.