1-1-2014

Changes In Cerebral White Matter, Vascular Risk And Cognition Across The Adult Lifespan

Andrew Robert Bender
Wayne State University,

CHANGES IN CEREBRAL WHITE MATTER, VASCULAR RISK AND COGNITION ACROSS THE ADULT LIFESPAN

by

ANDREW R. BENDER

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2014

MAJOR: PSYCHOLOGY (Behavioral & Cognitive Neuroscience)

Approved by:

Advisor	**Date**

DEDICATION

To Jyll and Luca

"If we all did the things we are capable of, we would astound ourselves."
— Thomas A. Edison

"The scientist only imposes two things, namely truth and sincerity, imposes them upon himself and upon other scientists." — Erwin Schrödinger

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Naftali Raz for his enduring support and mentorship, as well as my dissertation committee members for their time, insight, and guidance. I have been exceptionally fortunate to have so many positive influences and role models, as a scientist, father, and friend. I am overwhelming grateful for my friends, colleagues, and faculty mentors at the Wayne State University Institute of Gerontology and Department of Psychology for their mentorship, warmth and community these past several years. I wish thank Paolo Ghisletta, Boris Baltes, and Ana Daugherty, for their time and insights with statistical analysis for this project, and Cheryl Dahle and the members of the Raz lab, past and present, for their assistance in data collection.

To my parents, thank you for instilling in me a sense of curiosity and wonder, and emphasizing the importance of education, and love for learning. Finally, I wish to thank my wife Jyll for her unwavering love and support in my academic pursuits, for being my companion in life's adventures over the past decade, and for simply helping me be better in all ways.

This work has been possible via funding through the National Institute on Aging grant R37-AG11230 to Dr. Raz, the NIA/NIH training grant T-32 AG00275-06 to Peter Lichtenberg and the Institute of Gerontology, and the Wayne State University Thomas C. Rumble graduate fellowship.

TABLE OF CONTENTS

Dedication …… ii

Acknowledgments ……... iii

List of Tables ……..……...… v

List of Figures ……..…… vi

Chapter 1 – Introduction ……..…… 1

Chapter 2 – Methods ..…….……..…….. 16

Chapter 3 – Results ……...……..……... 43

Chapter 4 – Discussion …….……. 69

Appendix A: Spaghetti Plots of change in DTI indices across regions……..…… 78

References…….. 88

Abstract…..…….……..... 110

Autobiographical Statement……..... 112

LIST OF TABLES

Table 1: Participant characteristics ……..…… 17

Table 2: Univariate model results without covariates - DA …… 44

Table 3: Univariate model results without covariates - DR …… 45

Table 4: Univariate model results without covariates - FA …… 47

Table 5: Univariate model results without covariates - MD ……... 49

Table 6: Significant covariate effects on latent difference score factor …….. 50

Table 7: Comparison of cross-sectional correlations with age and mean differences …..... 52

Table 8: Goodness-of-fit for cognitive CFAs ……...……..…….... 54

Table 9: Significant baseline associations: DTI-Gf Letter Sets ……...……... 61

Table 10: Significant Paths in DTI-Gf Letter Sets Multivariate LDMs …….…… 62

Table 11: Significant baseline Associations: DTI-Names …….…. 63

Table 12: Significant Paths in DTI-EM (Names) Multivariate LDMs ...……..... 64

Table 13: Significant Paths in DTI-EM (Word pair task) Multivariate LDMs……... 66

Table 14: Significant baseline Associations: DTI-Word pair ….……..... 67

LIST OF FIGURES

Figure 1: WMH/CSF Masking Procedure.……...… 23

Figure 2: Univariate LDM Measurement Model ..……...… 38

CHAPTER 1

INTRODUCTION

Reduced integrity of cerebral white matter (WM) and elevations in vascular risk (VR) factors are implicated in age-related cognitive decline (Burgmans et al., 2010; Gunning-Dixon & Raz, 2000; Kennedy & Raz, 2009a; Madden et al., 2009, 2011). In the past decade, numerous studies investigating the neuroanatomical correlates of cognitive aging have employed diffusion tensor imaging (DTI) for *in vivo* quantitation of cerebral WM. However, few of the studies investigating associations between age-related cognitive decline and reduced WM integrity using DTI have evaluated *change* in cerebral WM. Although cross-sectional data can provide an informative snapshot of age-related differences, only longitudinal studies can elucidate individual differences in change over time or modifying factors (Lindenberger et al., 2011; Maxwell & Cole, 2007).

The limited reports comparing longitudinal change in WM with concomitant change in cognitive performance (Barrick, Charlton, Clark, & Markus, 2010; Charlton, Schiavone, Barrick, Morris, & Markus, 2010; Teipel et al., 2010) have several limitations: none have accounted for the effects of VR factors, all have used linear models that do not do assess heterogeneity in age-related change (see Raz et al., 2005, 2008), and all offer only limited anatomical specificity of DTI-cognition associations. In addition, these studies have included samples of middle-aged and older adults with mixed or unclear vascular pathology or dementia, rather than a well-controlled and characterized sample of healthy adults.

The present study addresses these questions in the context of existing data from a well characterized cohort of healthy participants covering the adult lifespan sample, measured on

two occasions, roughly two years apart. Such a project imposes numerous challenges, particularly regarding DTI image processing and data analysis. Specific concerns include registration and sampling of longitudinal DTI data, treatment of white matter hyperintensities (WMH), treatment of vascular risk factors, and statistical framework suitable for modeling heterogeneity in change for DTI, VR, and cognitive factors and their interrelationships.

Measurement of WM integrity

Volumetric assessments of normal appearing and lesioned WM have provided gross measures of WM integrity (Gunning-Dixon & Raz, 2000; Gunning-Dixon et al., 2009; Raz et al., 2005, 2010). Also referred to as leukoaraiosis, WMH are believed to reflect chronic vascular pathology (Brown & Thore, 2011; Burgmans et al., 2010; Dufouil et al., 2001; Gons et al., 2010; Kennedy & Raz, 2009b; Pantoni & Garcia, 1997; Raz et al., 2007; Raz et al., 2011b; Raz et al., 2012). However, WM volumetry provides a macroscopic index that is less sensitive to the effects of age and pathology than measures that take advantage of the microscopic properties of WM, such as the diffusion of water (Fjell et al., 2008; Hugenschmidt et al., 2008; Basser & Pierpaoli, 1998; Pierpaoli et al., 2001; Pierpaoli & Basser, 1996). These neuroimaging methods developed over the past two decades allow *in vivo* characterization and quantitation of microstructural properties of cerebral WM. DTI indices have been shown to be both more sensitive than WM volume to child and adolescent development (Westlye et al., 2009) and more strongly associated with adult aging (Burgmans et al., 2010; Fjell et al., 2008; Giorgio et al., 2010).

DTI

DTI quantifies the directionality and magnitude of the tissue-specific constraint of

molecular (Brownian) motion of water. In principle, whereas unconstrained water diffuses isotropically, hydrophobic myelin and axonal membranes in WM restrict water to diffuse preferentially along the length of myelinated axonal white matter fiber tracts (Beaulieu, 2002; Beaulieu & Allen, 1994; Concha, Livy, Beaulieu, Wheatley, & Gross, 2010). Thus, differential measures of diffusion and anisotropy, or the constrained non-random movement of water, can provide useful information about WM microstructure.

DTI studies most commonly utilize axial, diffusion-weighted, echo-planar imaging (EPI) sequences (Chen & Hindmarsh, 2001; Turner, Le Bihan, & Chesnicks, 1991). In order to calculate the tensor, diffusion must be independently sampled in at least six non-collinear directions (Moseley, Bammer, & Illes, 2002). Acquisition of these diffusion-weighted images requires a diffusion gradient with sufficient strength and pulse duration to sample diffusion in WM. The diffusion gradient is frequently reported as a high "b-value," commonly 1000 s/mm^2, on two to three axes; an image with no gradient weighting applied, or b_0 image is also required. From these separate images, the diffusion tensor, a geometric expression that describes a three-dimensional ellipsoid (Basser, Mattiello, & LeBihan, 1994) is calculated using a 3×3 matrix constructed from the different diffusion-sensitized directional images. Using matrix diagonalization, three eigenvectors are calculated which describe the directionality of diffusion along the three principal axes (x, y, z), and corresponding eigenvalues which provide indices of magnitude of diffusion (see Kingsley, 2006 for a complete description of DTI mathematics). Together, the three eigenvalues λ_1, λ_2, and λ_3, ($\lambda_1 \geq \lambda_2 \geq \lambda_3$) describe the shape of the ellipsoid. Mean diffusivity (MD) is calculated as the average of the three eigenvalues, and is mathematically equivalent to the apparent diffusion

coefficient (ADC) reported in diffusion-weighted imaging (DWI).

The first eigenvalue λ_1 provides the principal diffusional direction and is also used as an index of axial diffusivity (DA). Similarly, averaging λ_2 and λ_3 yields a mean estimate of diffusion perpendicular to λ_1, or radial diffusivity (DR or $\lambda_\perp$). Unconstrained diffusion is isotropic, whereas myelin, cellular membranes, and fiber packing density restrict diffusion (Beaulieu, 2002; Beaulieu et al., 1996; Concha et al., 2010). This is also reflected in the shape of the tensor ellipsoid with cigar-shaped, prolate ellipsoid reflecting highly constrained diffusion, and uniform or oblate spheroids indicative of more isotropic diffusion. Thus, indices of anisotropy provide information about the restricted movement of water, which corresponds to microstructural properties of WM, particularly the directional coherence of WM fibers within a voxel. Fractional anisotropy (FA) is a scalar index that describes the eccentricity, or deviations from sphericity in the diffusion ellipsoid (Basser & Jones, 2002); in other words, FA provides an index of the magnitude of constraint or non-randomness in diffusion. FA has been used as a marker of intra-voxel coherence or fiber alignment consistency. As an index of intra-voxel coherence, FA is reduced in many regions that include an admixture of multiple, differentially oriented WM fiber tracts (e.g., centrum semiovale, occipital forceps, posterior cingulate, precuneus, arcuate fasciculus). Conversely, FA is maximal in regions where fibers share a common alignment, such as the splenium of the CC. Thus, in voxels with crossing fibers and reduced FA, the tensor shape may not simply be oblate or prolate, but vary along other dimensions.

DTI in aging

The most commonly reported DTI indices in studies of aging and cognition include

MD, FA, DA, and DR, which appear to reflect different tissue properties (Beaulieu, 2002); however, most published DTI studies of cognitive aging do not report all such indices. In addition, different patterns of these values are associated with different markers of pathology. Increased DR, accompanied by reduced DA may be indicative of Wallerian degeneration (Ford & Hackney, 1997; Ford, Hackney, Lavi, Phillips, & Patel, 1998; Pierpaoli et al., 2001; Sun et al., 2008). More specifically, ex vivo animal studies have shown myelin breakdown is associated with increased DR, whereas reduced DA appears to reflect axonal damage (Song et al., 2003; Song et al., 2002; Sun et al., 2008; Sun et al., 2006). Thus, Wallerian degeneration in isolated fiber bundles is associated with reduced DA and increased DR (Pierpaoli et al., 2001). In addition, FA appears influenced by myelin, but is largely dependent on fiber packing density, axon diameter, axonal membranes and intercellular space (Beaulieu, 2002). Findings from *ex vivo* DTI and histology of spinal cords from patients with multiple sclerosis (MS) also supported increased DR as reflecting demyelination and axonal loss, except in lesioned areas with marked alterations in WM architecture (Klawiter et al., 2011). Moreover, the two diffusivity measures may differently reflect timing of insult, with reductions in DA following acute injury, whereas DR increases are manifest over a longer period of time (Klawiter et al., 2011; Naismith et al., 2009; Pitkonen et al., 2012; Song et al., 2003; Song et al., 2002; Sun et al., 2008; Sun et al., 2006; Zhang et al., 2009). However, findings from *ex vivo* studies should be interpreted with some caution as formalin fixation reduces diffusivity measures, but not anisotropy in both human and animal models (Schmierer et al., 2008; Sun, Neil, & Song, 2003). Moreover, all of these reports are inherently cross-sectional and do not inform about change over time within the individual. An

additional limitation is these findings are predominantly restricted to WM tracts with high coherence, and not in regions in which DTI signal may be complicated by the presence of differential directions in the fiber populations.

Aging in healthy adults is associated with reduced FA and increased MD, possibly due to increased DR (Bhagat & Beaulieu, 2004) associated with myelin and fiber loss. However, patterns differ across studies and tasks and much speculation has been made regarding the underlying neuroanatomical interpretation of such effects (see Madden et al., 2011 for a review). More accurate interpretation of different combinations of DR, and DA requires better understanding of other key influences on λ_1, λ_2, and λ_3 within a voxel, such as the presence of crossing fibers (Vos, Jones, Jeurissen, Viergever, & Leemans, 2011a). Madden et al. (2011) evaluated reports of correspondence between cognition and FA because that is the most commonly reported DTI metric. However, any claims regarding specific anatomical attributes underlying the DTI metrics require more than a single tensor model. That is, suggestions that DR and DA specifically reflect myelination or axonal integrity require additional WM measures more specific to myelination or greater geometric specificity (Wheeler-Kingshott & Cercignani, 2009; Jones et al., 2013). Thus, controlling for such influences (Douaud et al., 2011) may bolster inferential speculation about spatial patterns of DR and DA, and the correspondence with underlying cellular or pathological processes associated with aging (Bennett, Madden, Vaidya, Howard, & Howard, 2010a; Burzynska et al., 2010; Lebel et al., 2012) or learning (Engvig et al., 2011).

Aging and WM tracts

Cross-sectional evidence points to heterochronous patterns of WM development and

decline across fiber tracts (Lebel & Beaulieu, 2011). For association tracts, however, there is a general pattern of marked increases in myelinated fibers and WM integrity throughout childhood (Taki et al., 2012), followed by a more shallow, positive slope or plateau from the third decade into middle age, and subsequent accelerated decline from the fifth or sixth decade on (Hasan et al., 2009a; Hasan et al., 2009b; Kochunov et al., 2012; Lebel & Beaulieu, 2011; Lebel et al., 2012). However, this does vary with some tracts such as the cingulum bundle exhibiting a far shallower negative slope in adulthood in comparison to inferior frontal-occipital fasciculus, which has a far steeper trajectory after the fourth decade (Lebel et al., 2012). In addition, findings from longitudinal DTI studies in healthy adults are very limited. To date, only a handful of studies have looked at the natural course of change over time in DTI metrics in adults (Barrick, Charlton, Clark, & Markus, 2010; Charlton, Schiavone, Barrick, Morris, & Markus, 2010; Sullivan, Rohlfing, & Pfefferbaum, 2010a; Teipel et al., 2010). Fiber tracking data from healthy adults showed greater FA reductions in CC body and genu than the splenium (Teipel et al., 2010). Using a tract-based approach to evaluate longitudinal DTI data, Barrick et al. (2010) reported that despite a lack of accelerated WM decline in older age and inconsistent longitudinal changes within a tract, that cross-sectional analysis of baseline data underestimated longitudinal decline. Moreover, the only study to evaluate associations between longitudinal changes in WM integrity and cognition due to normal aging (Charlton et al., 2009) employed whole brain histograms of FA and MD and sheds little light on anatomically specific WM correlates of cognitive aging.

DTI & cognitive aging

The literature investigating relationships between WM integrity and variability in

cognitive performance using DTI is predominantly cross-sectional. Moreover, comparison of cross-sectional findings reveals a lack of stability of such associations across studies (see Madden et al., 2009, 2011 for reviews). However, the extant literature does show strong correspondence between cerebral WM integrity and cognitive processing speed. For example, numerous reports suggest lower WM integrity is associated with slower processing speed in healthy adults (Burgmans et al., 2011; Correia et al., 2008; de Groot et al., 2000; Gold, Powell, Xuan, Jiang, & Hardy, 2007; Kennedy & Raz, 2009a; Kochunov et al., 2012; Liston et al., 2006; Liu et al., 2011; Madden et al., 2004; Penke et al., 2010b; Stebbins et al., 2001b; Tuch et al., 2005; Turken et al., 2008). Thus, evaluation of DTI metrics of WM integrity can provide important insights into the neurobiological bases of age-related slowing and associated decrements in other cognitive, perceptual, and motoric abilities (Eckert, 2011).

Connectivity and anatomical association between prefrontal cortices and medial temporal and parietal cortices are implicated in performance on measures of episodic memory (Buckner et al., 1999; Bucur et al., 2008; Grady, McIntosh, & Craik, 2003; Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006; Kramer et al., 2005; Metzler-Baddeley, Jones, Belaroussi, Aggleton, & O'Sullivan, 2011; Swick & Knight, 1999) as well as executive functioning and working memory (Della-Maggiore et al., 2000; Kane & Engle, 2002; Kennedy & Raz, 2009a; Madden et al., 2007; Miller & Cohen, 2001; Ziegler et al., 2010). Age-related cognitive decrements are associated with reduced structural and functional connectivity between prefrontal cortices and other regions (Andrews-Hanna et al., 2007; Clapp, Rubens, Sabharwal, & Gazzaley, 2011; Grady et al., 2003). Early DTI studies of cognitive aging reported reduced anterior FA in older adults (Pfefferbaum et al., 2005; Salat

et al., 2005) was associated with reduced performance on measures of episodic memory, working memory, speed of processing, and reasoning (Madden et al., 2004; Stebbins et al., 2001a; Stebbins et al., 2001b).

Findings from functional and anatomical studies of cognitive aging have supported disconnection hypotheses in explaining age-related decrements in cognitive function (Geschwind, 1965). However, the regions and their connective substrates vary by cognitive domain. More specifically, anterior-posterior dysconnection may be implicated in reduced working memory (Davis et al., 2009), whereas impairments in frontal WM may be associated with reduced executive functioning (Buckner et al., 1999; Bucur et al., 2008; Grady et al., 2003). Other cross-sectional findings from studies of healthy adults and patients with lesions have shown frontal-temporal disconnection may underlie episodic memory impairments by way of uncinate fasciculus (Papagno et al., 2010), cingulum bundle and fornix (Metzler-Baddeley, 2011; Sasson et al., 2010, 2011).

Dearth of relevant prior findings

Given the dearth of longitudinal studies of DTI and aging, forming specific hypotheses regarding relationships between change in DTI indices (e.g., FA and MD) and change in cognition over a two-year delay is challenging. Moreover, the extant longitudinal DTI aging studies limited their analyses to either specific fiber tracts estimated with various tractography methods (Sullivan et al., 2010), or used both voxel-wise and region of interest (ROI)-based repeated measures analysis without accounting for additional continuous covariates such as age (Teipel et al., 2010). Similarly, other longitudinal approaches for evaluating age-related changes in DTI-based indices and cognition limited analyses to whole

brain or slice-by-slice histograms of FA and MD (Barrick et al., 2010; Charlton et al., 2010). Most of those studies employed suboptimal statistical techniques for assessing change, such as difference scores, paired t-tests or repeated measures ANOVA (Bereiter, 1963; Francis, Fletcher, Stuebing, Davidson, & Thompson, 1991; Overall & Woodward, 1975). As DTI has unclear test-retest reliability, the consequence of inclusion of measurement error in longitudinal assessment using difference images is uncertain. Furthermore, as cross-sectional findings cannot inform about intraindividual change or variability therein (Lindenberger et al., 2011), even comparison of two occasions provides a better index of change.

Need for longitudinal data

Unfortunately, it is possible that the increasing number of cross-sectional findings may be altogether missing the mark regarding individual differences in change of WM (Lindenberger et al., 2011). One reason may be that heterogeneously manifest VR factors may yield longitudinal effects that cannot be characterized by cross-sectional data. Although diagnosed hypertension may be associated with a steeper 5-year decline in volumes of some brain regions (Raz et al., 2005), preliminary evidence shows an altogether different pattern for DTI indices. Bender, Daugherty, & Raz (2012) evaluated the effects of age and hypertension on change in FA and MD over 15 months. They reported that in normotensive healthy adults, MD increased and FA decreased across multiple regions, including cingulum bundle, CC, and fornix. In contrast, individuals who are medically treated for essential hypertension showed the opposite pattern. That is, rather than showing exacerbated declines in WM integrity over two occasions, hypertensive participants who reported receiving medical treatment showed improvements. One possible interpretation suggests that the effects

of sub-clinical hypertension and other nascent VR factors may damage WM integrity in healthy persons. Thus, longitudinal analysis is essential to elucidating the real patterns and modifiers of change, even over a short delay.

Statistical concerns

Although more easily implemented in DTI analysis software, standard linear statistical models such as paired t-tests and repeated measures variance partitioning methods are problematic for evaluation of ongoing processes (see Raz et al., 2005, 2008). In particular, such methods are unable to model individual differences or the impact of factors contributing to heterogeneity in change. This limitation may be especially pronounced when assessing change over a short period where error variance may be larger than longitudinal effects. That is, decline or improvement between two closely spaced measurement occasions may be more suspect as it is unclear whether observed change may be due to sampling variation. This demonstrates a clear need for latent modeling approaches, although this imposes other concerns regarding sample size and statistical power which is often lower in prospective than cross-sectional studies due to attrition.

In addition to measurement of change in DTI, there are also concerns regarding relating change in WM with change in cognition over a limited period. Specifically, in order model the influence of other factors on longitudinal change, such change must demonstrate significant variance. Conversely, if no such significant individual differences are present, then there is no variability in change to predict. Because individual differences in change may be attenuated in latent modeling approaches, it is possible that some cognitive domains may not show significant variance in change in a healthy sample. Furthermore, if latent changes

are observed in cognitive measures, it is possible that they reflect retest effects, or performance improvements following repeated administration of testing materials. Such retest effects are particularly pronounced for mnestic measures and may obscure age-related decline (Ferrer, Salthouse, Stewart, & Schwartz, 2004; Salthouse, Schroeder, & Ferrer, 2004). Thus, mean improvement on cognitive measures as assessed with latent modeling approaches likely reflects retest effects, rather than age-related change or decline.

Challenges of longitudinal DTI

In addition to concerns regarding statistical modeling of change between two measurement occasions, there are issues specific to longitudinal measurement of DTI indices. First, greater slice thickness and variability between MRI measurements in slice placement may reduce measurement reliability between occasions. However, spatial normalization of DTI maps may result in increased interpolation and altered signals (Chao, Chou, Yang, Chung, & Wu, 2009). Thus, particular care must be taken to both maximize spatial correspondence between measurements and minimize interpolation of the DTI signal. Diffeomorphic registration methods can be used to register images between measurements, and thereby minimize these issues as long as data from all measurements is treated identically (Engvig et al., 2011; Huang et al., 2012; Reuter et al., 2012).

Longitudinal DTI data may have been originally collected using imaging standards that are no longer considered optimal. Thus, one concern is how to make the most of the data in light of such limitations. Although acquisition of six gradient directions is minimally sufficient for calculation of the diffusion tensor, it is neither optimal for improving signal-to-noise ratio for FA or MD nor sufficient for DTI tractography (Jones & Cercignani,

2010; Mukherjee, 2008). Increasing the number of gradient directions would improve SNR more than increasing the number of averages. However, Danielian and colleagues (2010) showed that increasing the number of averaged scans improves longitudinal reliability for tractography. With multiple averages concatenated at the scanner console, this could be viewed as problematic given the lack of specific registration across averages, particularly for voxel-wise comparison. However, it might also be seen positively as similar to applying a smoothing function to ameliorate the effects of minor movement or differences between averages or occasions (Engvig et al., 2011) resulting in a more generalizable signal. Regardless, even when the original DTI sequence parameters are suboptimal in comparison to more modern standards (e.g., six gradient directions vs. 32), longitudinal data can provide information that even the most optimal cross-sectional measures cannot, particularly when all possible measures are taken to maximize spatial correspondence and minimize any undue or differential interpolation between measurements.

Vascular Risk

In general, the influence of vascular risk on the relationship between DTI measures of WM integrity and cognitive abilities has not been addressed in a well controlled and characterized sample using statistical methods best suited to such analysis. Few studies have investigated the influences of aging and VR on the associations between DTI indices and cognitive abilities. In some cases, this was limited to ad hoc comparisons between hypertensives and normotensives on cognitive measures (Kennedy & Raz, 2009a) or DTI metrics (Sasson, Doniger, Pasternak, & Assaf, 2010). Others modeled BP, HTN or other VR indices as specific variables of interest (Hannesdottir et al., 2009; Leritz et al., 2010; Vernooij

et al., 2009). Unfortunately, the growing acceptance that a vascular and inflammatory etiology underlies, at least in part, neurocognitive aging and dementia (Franceschi et al., 2007; Fung, Vizcaychipi, Lloyd, Wan, & Ma, 2012; Piccinin, Muniz, Sparks, & Bontempo, 2011; Schneider & Bennett, 2010; Ungvari et al., 2010; Wright & Sacco, 2010) appears largely unacknowledged by cognitive aging studies using DTI that are not specifically investigating the modifying roles of VR.

Most studies neither exclude nor account for the presence of individuals with common VR factors (Sasson et al., 2011). Although numerous studies estimate age-related variance in DTI indices (Charlton et al 2006; see Madden et al 2011 for a review), accounting for individual differences in VR factors may be at least as informative (Vernooij et al., 2008; Vernooij et al., 2009). Furthermore, in light of age-related increases in prevalence for such factors (Ervin, 2009), some proportion of the negative effects of age on WM integrity reported by many studies may instead reflect an admixture of age and unidentified VR factors. Conversely, those studies that control for the effects of age may also be partialing out some proportion of variance conveyed by VR that may be either dissociable from or collinear with age. In addition, VR-associated variance in DTI metrics in HTN and pathology may be qualitatively different from that of healthy aging, due in part to differences in underlying vascular pathology.

Significance of the study

The present study is intended to address prescribed gaps in the extant literature relating age, DTI indices of cerebral WM, and cognition while controlling for the effects of hypertension. Moreover, use of latent difference score modeling (LDM; McArdle &

Nesselroade, 1994) permits estimation of longitudinal effects on the level of latent variables (LVs), free of measurement error. Importantly, this approach allows separate modeling of variance in LVs, and thereby permits examination of influences among individual differences in both individual LVs and in change between occasions of measurement. In addition, this study represents the first longitudinal DTI study of healthy participants covering the adult lifespan representing a 'best case scenario' of optimal aging. Given the greater theoretical value of longitudinal data in elucidating change, we sought to demonstrate the utility of sampling change over a relatively short period, even using suboptimal methods for DTI acquisition.

We expected that the LDM framework is a feasible approach for assessing change in DTI data, and that meaningful changes in DTI indices are observable over 2 years. In addition, we hypothesized these latent differences vary across brain regions; moreover, we expected that the extant cross-sectional findings would not provide an a priori basis for where to look for change, and even the nature of the longitudinal differences. However, we expected that different DTI indices should be differentially representative of underlying of microstructural properties of cerebral WM, although exactly which properties, we could not be sure of. We did expect that variance in change in DTI measures would be associated with variance in cognition.

CHAPTER 2

METHODS

Participants

The sample was drawn from data collected as part of the ongoing longitudinal study of brain and cognitive aging conducted by Dr. Naftali Raz and the Cognitive Neuroscience of Aging Laboratory at Wayne State University. A total of 104 participants were scanned on the Bruker Biospin 4T MRI scanner during the first two waves of the cohort whose initial measurement was collected from June 2005 to February 2009. All participants completed a thorough self-report health questionnaire

All participants were screened via self-report questionnaire to rule out depressed state (CES-D; Radloff, 1977; cut-off = 15), and an experimenter administered the Mini Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975; cut-off = 26) to screen for cognitive impairment. The experimenters screened all participants for near, far, and color vision problems (Optec 2000 Vision Tester, Stereo Optical Co., Inc., Chicago, IL) and speech-range hearing deficits (MA27 Screening Audiometer, Maico Diagnostics, Eden Prairie, MN).

Eight out of the 104 participants who underwent MRI at both occasions on the 4T magnet were subsequently excluded from analysis. These included: a 41-year old normotensive, Caucasian female for incidental finding of arteriovenus malformation; a 71-year old, hypertensive, Caucasian female for MMSE score of 25 at follow up assessment; a 64-year old, normotensive, Caucasian female who was missing part of baseline DTI data; a 39-year old, normotensive, Caucasian female, excluded for prior surgical repair of a cerebral

aneurism not reported until follow up assessment; a 52-year old, normotensive, black female, for high fasting blood glucose at both baseline and follow up (165 mg/dL & 192 mg/dL, respectively). In addition, we excluded three participants for depressed state as indicated by CES-D scores over cut-off at follow up assessment. These included a 55-year old, normotensive, Caucasian female (CES-D = 18), a 63-year old, hypertensive, black male (CES-D = 25), and a 31-year old, normotensive, Caucasian female.

The final sample consisted of 96 participants including 66 women and 30 men, ranging in age from 17 to 78 years at baseline assessment (Table 1). Men and women did not differ with regard to mean age, MMSE scores, self-reported years of education, self-reported engagement in regular exercise and frequency of exercise, or body mass index (BMI). However, men had significantly higher diastolic blood pressure and proportion of treated hypertensives, and marginally higher systolic blood pressure than women. In addition, women had one month longer delay on average between MRI scans than men.

In addition to both occasions of DTI data, all participants had at least one occasion of each cognitive and VR measurement. Assessment of blood pressure and hypertension is described below under *Measures of vascular risk*.

Table 1. Participant characteristics

	Women	Men		
Variable	Mean (SD)	Mean (SD)	t or χ^2	p
Age (years)	54.59 (13.44)	55.3 (14.12)	-0.236	.814
Delay (months)	25.44 (2.13)	24.54 (2.12)	1.932	.056
MMSE	29.03 (0.99)	28.87 (0.97)	0.754	.453
Education	15.73 (2.23)	15.47 (2.58)	0.505	.615
Systolic BP	120.24 (12.75)	125.41 (11.2)	-1.912	.059

Diastolic BP	73.86 (6.64)	77.81 (8.56)	-2.461	.016
% Exercise	78.8%	83.3%	0.268	.604
Days Exercise	3.17 (2.14)	3.83 (2.19)	-1.406	.163
BMI	3.26 (0.21)	3.33 (0.16)	-1.689	.095
% HBP Dx	15.15%	30.00%	4.134	.042

Notes: BP=blood pressure; BMI=body mass index; Dx=diagnosis.

'Missingness' in longitudinal data

A total of 219 participants (67.6% women) completed baseline assessments. Of those, 134 completed follow up assessments. The 134 participants who completed both waves of the study, did not differ from the 85 who did not return with regard to proportion of men and women ($\chi^2 = 2.600$, $p = 0.107$), frequency of physician-diagnosed and treated hypertension ($\chi^2 = 1.204$, $p = 0.273$), number of years education ($t[217] = 0.275$, $p = 0.784$), systolic ($t[216] = 0.329$, $p = 0.743$) and diastolic blood pressure ($t[216] = 1.28$, $p = 0.204$), self-reported exercise ($\chi^2 = 0.881$, $p = 0.348$), and frequency thereof ($t[217] = 1.651$, $p = 0.100$), smoking ($\chi^2 = 1.003$, $p = 0.317$). However, the mean age and MMSE scores were significantly lower among those who did not return (mean age = 48.49, SD = 17.32 years; mean MMSE = 28.60, SD = 1.18) in comparison to participants completing both waves (mean age = 53.37, SD = 13.82 years; mean MMSE = 28.98, SD = 0.99; $t[217] = -2.304$, $p < .05$, and $t[217] = -2.551$, $p < .05$, respectively).

Moreover, of the 134 participants who completed follow up testing, 104 were scanned on the Bruker 4T scanner, but due to hardware upgrades, the remaining 30 were scanned on a Siemens Verio 3T magnet. Differences in magnet strength precluded comparison with baseline measurements. The 96 participants included in analysis (mean age = 54.698, SD = 13.719 years) with MRI data at both measurement occasions, were older than the 30

participants subsequently scanned on the 3T (mean age = 48.774, SD = 13.987 years; t[125] = 2.080, p = 0.040), but did not differ with regard to proportion of men and women, self-reported years education, presence or frequency of exercise, smoking, frequency of hypertension, MMSE score, or systolic and diastolic blood pressures (p > .100 for all).

MRI imaging

MRI images were originally acquired as part of a longer protocol, using a Bruker MedSpec 4T scanner equipped with an 8-channel head coil. A 1 slice 2D localizer was initially acquired in the sagittal plane using the following parameters: Voxel size: 2.2 mm × 2.2 mm; 10 mm slice thickness; FOV = 280 mm; TR = 20 ms; TE = 5 ms. A 2-D echo planar diffusion-weighted sequence acquired images with the following parameters: TR = 4900 ms; TE = 79 ms; 41 slices; slice thickness = 3 mm; distance factor = 0; FOV = 256 mm; matrix = 128 × 128; voxel size = 2.0 mm × 2.0 mm × 3.0 mm; GRAPPA acceleration factor = 2. Diffusion weighted data were collected in six orthogonal gradient directions using a diffusion weighting of 800 s/mm^2; an additional T2-weighted image was collected without diffusion weighting (b_0 = 0 s/mm^2).

DTI processing.

Images were averaged across acquisitions on the console and written into DICOM format. We converted image data from DICOM format into NIFTI-1 (.nii) using MRIConvert 2.0 (Jolinda Smith, University of Oregon, Lewis Center for Neuroimaging [http://lcni.uoregon.edu/~jolinda/mriconvert]). During conversion, the software extracted and rotated the b-vector matrix for each participant by multiplying the original b-vector gradient matrices by the inverse of the individual rotation matrices (i.e., the

patient image orientation matrix). This is preferable to using a global b-vector matrix based on scanner acquisition as it improves directional accuracy of the calculated tensor parameters.

We developed a custom DTI processing pipeline, written in a Bourne shell (bash) script, using tools freely available in the FMRIB Software Library (FSL) v5.0.2 (Analysis Group, FMRIB, Oxford, UK). The pipeline was written, tested and run on an Apple MacPro workstation running Mac OS 10.7 (Apple, Inc., Cupertino, CA). Pipeline functionality included DTI preprocessing, tensor fitting, segmentation of WMH and CSF, and optimized registration of longitudinal pairs, as well as nonlinear deprojection of a group-wise WM skeleton and atlas labels and data sampling. The pipeline was designed to optimize registration between measurements within each participant for generating the WM skeleton, while minimizing possible interpolation resulting from diffeomorphic transformation of DTI data. The pipeline combined traditional DTI pre-processing, diffeomorphic registration, tract-based spatial statistics (TBSS; Smith et al., 2006) procedures for skeletonisation, and probabilistic WM atlases to generate regions of interest (ROIs) on the WM skeleton. Registration procedures were adapted from several sources (Amlien et al., 2012; Engvig et al., 2012; Huang et al., 2012) with the overarching goal of optimizing anatomical correspondence between each participant's two measurements, prior to skeletonisation. The following section details specific steps and tools used in DTI processing.

*DTI pre-process*ing. We extracted copies of the first (b_0) volume of each of the two paired (Time1, Time 2) 4D nii files. Next, the FMRIB's Linear Image Registration Tool (FLIRT) linearly registered the b_0 images from each occasion using a 6-degrees of freedom (df) rigid registration with tri-linear interpolation, and the transformation matrices were saved

for subsequent use. Following registration, the script invoked the 'avscale' command to calculate the matrices needed to transform both images into the intermediate or 'halfway' space between them. In other words, although initial registration treated the baseline image as the target to which the Time 2 image was registered, this information could be used to determine secondary transformation matrices; these new transformation matrices could then be used to register each image to an intermediate space. The script called the 'convert_xfm' command to extract the halfway-forward and halfway-backward matrices and invert copies of the two matrices for later use.

We used the brain extraction tool (BET; Smith, 2002) in FSL on both the paired, untransformed b_0 images and the native 4D .nii file to produce brain masks and remove non-brain tissue; BET parameters included the –m and –R flags for binary mask creation and robust, repeated estimation of the brain center, respectively. The b_0 images, stripped of non-brain tissue were then eroded by one voxel using the fslmaths –ero function. We then used the 4D brain masks to fit the tensor using the dtifit function, as well as participant specific b-vectors, a common b-value of 800 mm^2/s, and the 'save_tensor' function to retain the DTI components for later use. The script called the fslmaths routine to average the second and third eigenvalues created by dtifit into radial diffusivity (DR) images for both occasions, and the first eigenvalue map was now designated as the axial diffusivity (DA) image.

Next, the pipeline invoked the 'vecreg' command to rotate the saved tensor components using the halfway transformation matrices (i.e., T1 – halfway backward, Time 2 – halfway forward). The fslmaths 'tensor-decomp' function refit the tensor data for both occasions in the transformed, halfway space between the two longitudinal images. We used

fslmaths to apply an upper threshold of 1.2 to the resulting, halfway-transformed FA maps to remove any noisy voxels. These thresholded FA maps were subsequently used in the TBSS processing approach for WM skeletonisation, (not voxel-wise analyses or sampling).

Following tensor fitting and FA calculation, TBSS processing includes multiple steps. First, FA maps for the entire sample are non-linearly registered and individual FA maps are resampled to 1 mm^3 voxels. In the present study we used the FMRIB58 standard space FA image as the registration target. Next, FA maps for the entire sample are merged, and a mean FA image is created from this merged 4D file. This mean FA image was then skeletonised, and a 0.3 threshold for FA was applied in the present study. Last the individual FA maps were projected onto the skeletonised mean FA image, using peak intensity FA values from innermost voxels within each tract as a guide. These maps form the basis for skeletonised FA data in a common space and provide the spatial transformations needed to inversely warp or 'deproject' the data from standard to native space.

Preprocessing – WMH/CSF Segmentation. Next, the pipeline used FMRIB's Automated Segmentation Tool (FAST) to segment the eroded, skull-stripped b_0 images into 6 separate maps based on voxel intensity. These included two images representing primarily white matter (WM), in addition to images whose intensity reflects cerebrospinal fluid (CSF), areas of WM hyperintensities (WMH), or image noise at interfaces between CSF and other tissue types, GM, and hypointense voxels reflecting noise or near iron containing nuclei such as the basal ganglia. We then averaged and binarized the WM maps into a mask using fslmaths. Last, we refit the diffusion tensor using the same procedures as above, but using the binarized WM mask and re-calculated the radial and axial diffusion maps for the

WMH/CSF-masked data.

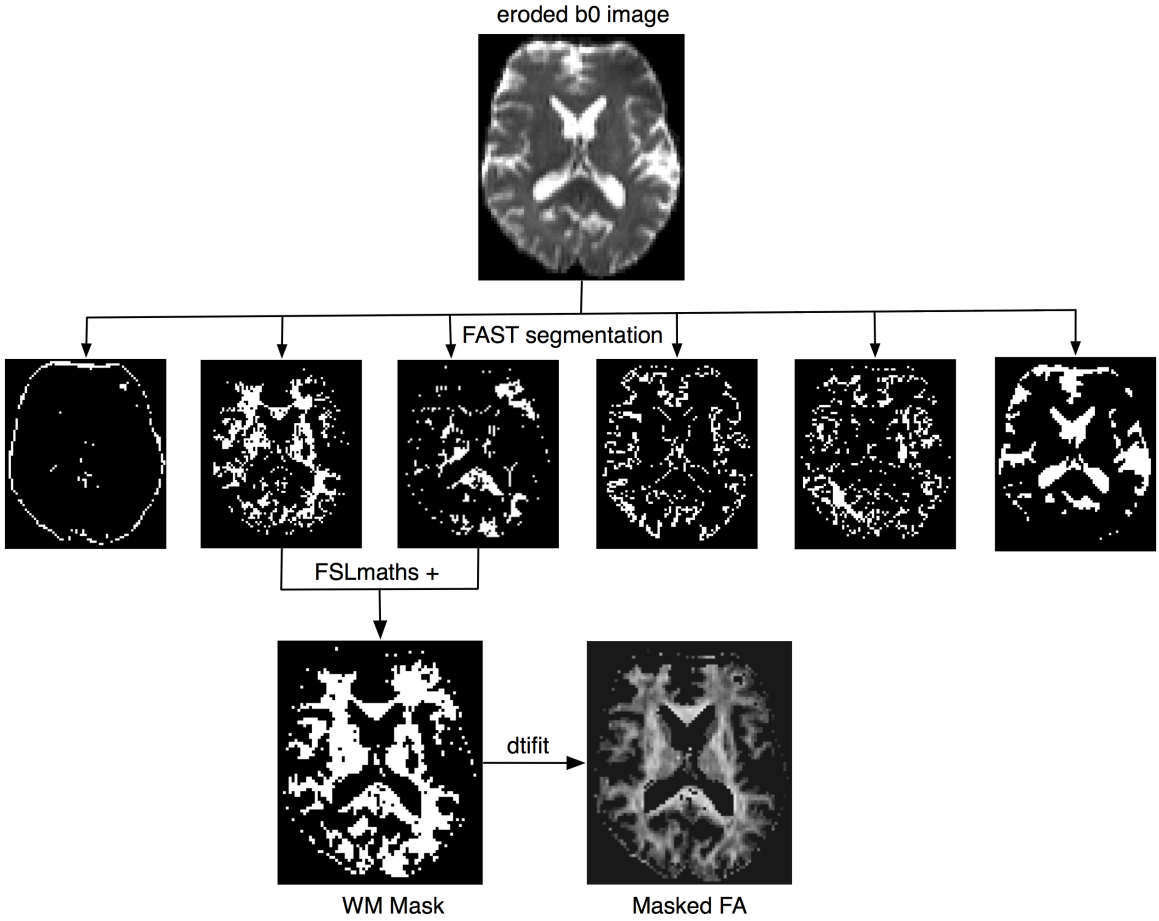


Figure 1. Diagram depicting the procedure used for segmenting out cerebral spinal fluid (CSF) and white matter hyperintensities.

TBSS-skeletonisation. We used the tract-based spatial statistics (TBSS; Smith et al., 2006) processing framework to create a group-wise WM skeleton in standard space, and then to nonlinearly deproject the skeleton and WM atlases back to halfway space. We ran standard TBSS processing on the FA images generated by refitting the tensor in halfway transformed space. The FMRIB58_FA standard space image was used as the target for non-linear registration, and the data were nonlinearly aligned into 1 mm × 1 mm × 1mm MNI 152 space. The TBSS pipeline then generated the mean, standard space FA image from both waves of the sample, and the corresponding WM skeleton. In the final step of TBSS processing (i.e.,

tbss_4_prestats), we applied a threshold of 0.3 to the mean WM skeleton, to reduce areas with poor reliability. Last, we ran the tbss_deproject routine on the mean WM skeleton mask, as well as to the JHU-ICBM white matter atlas labels at 1mm, and the JHU-ICBM white matter tractography atlas (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases); the '2' and '–n' flags were used to nonlinearly warp the skeleton mask and atlases back to the space they were in at the first step of TBSS processing – here, the halfway space between measurements, while maintaining the integer values of the atlas regions. Furthermore, by using tbss_deproject on the atlases, the atlas values were only deprojected along the WM skeleton.

Deprojection. We used FSL's FLIRT process to transform the atlas-derived, skeletonised ROIs and the mean WM skeleton mask from individual halfway space to the original, native space from acquisition. Using fslmaths, we applied a lower threshold of 0.20 and an upper threshold of 1.001 to the native FA images from both occasions. The thresholded FA image was then binarized and used to mask the deprojected skeleton mask in native space, reducing noise from each skeleton from which we sampled values. We used the FSL 'applywarp' function to deproject the Harvard-Oxford subcortical atlas (Desikan et al., 2006) to native space, using the subject-specific inverse warp matrices generated by TBSS and the halfway transformation matrices created earlier; whereas tbss_deproject restricts deprojected values to the WM skeleton, this deprojection method does not.

Mask creation. In a separate bash-scripted process, we used fslmaths to extract separate masks from the atlases deprojected to native space, for each subject. For redundancy, we used the individual native space, FA-thresholded skeletons as a secondary mask on the ROI masks during extraction. If possible, we extracted separate masks for left and right and combined

hemispheres.

In addition to mask extraction from the deprojected, skeletonised atlases, some additional manipulation of the masks was necessary. The atlases contained only individual masks for non-lateralized structures on the midline, such as corpus callosum (CC) and fornix. Therefore, we used fslmaths to extract separate hemispheric masks from the Harvard-Oxford atlas transformed to individual subject space. We used these to mask left and right sides of the CC including body (CC body), genu (CC genu) and splenium (CC splenium), to create separate lateralized masks. In addition, we sought to eliminate overlap between ROIs. The JHU-tractography atlas masks for forceps major (FMaj) and minor (FMin) were spatially redundant with the masks for CC genu and CC splenium. However, as we wished to model the variance from those regions separately we used fslmaths to create separate, non-overlapping masks by subtracting the masks for genu and splenium from those for forceps minor and major, respectively.

In addition, visual inspection of the uncinate fasciculus (UF) masks revealed substantial overlap between the two atlases, in this rather small ROI. Therefore, we used fslmaths to sum the separate masks from the two atlases into a new UF mask with increased coverage.

Data Sampling. Based on visual inspection of the resultant masks, we chose 14 WM atlas-derived ROIs for sampling and analysis. The WM masks included the following tracts and regions previously associated with cognitive abilities in studies of aging. These included CC genu, CC splenium and CC body, dorsal cingulum bundle (CBd), ventral cingulum bundle (CBv), superior longitudinal fasciculus (SLF), and three divisions of the internal

capsule, the anterior limb (ALIC), posterior limb (PLIC), and retrolenticular limb (RLIC) – all taken from the ICBM-DTI-81 white matter labels atlas (Mori et al., 2005; Wakana et al., 2007). In addition we included ROI masks taken from the JHU WM tractography atlas, including the inferior longitudinal fasciculus (ILF), inferior frontal-occipital fasciculus (IFOF), FMaj and FMin. The UF mask was averaged from the two atlases. Care was taken to ensure masks did not overlap. We chose to exclude masks that demonstrated visually apparent inconsistencies in coverage across the sample. These included the masks for superior frontal-occipital fasciculus, corticospinal tract, and corona radiata.

We then created a separate bash script to sample data from the atlas-derived masks. The script called the fslstats function to sample and output the mean values and standard deviations for non-zero voxels for masks from left, right, and combined hemispheres for all four DTI indices fractional anisotropy FA, mean diffusivity (MD), radial diffusivity (DR), and axial diffusivity (DA). The script sampled from both the non-masked data and from the data masked for WMH/CSF. We inspected standard deviations from the sampled FA values to help ensure no subject has excessive noise in a given ROI (e.g., standard deviations for all FA data were < .15). In addition, we evaluated the standard error of the mean values for each region sampled, for each participant. Standard error values were low (FA SE range = 0.0059 to range 0.08), and varied according to the size of the region sampled.

Measures of vascular risk (VR)

Diagnosis and treatment of hypertension. The number of years that participants with self-reported diagnosis of hypertension have been taking antihypertensive medication, as reported at follow up testing, served as a VR measure. This measure was chosen as a

covariate representing hypertension-related vascular risk for several reasons. First, whereas a dichotomous indicator of hypertension diagnosis contains little meaningful variance, duration of medication includes measurement variance that is likely indicative of the severity of effects. Moreover, it is also possible that those taking anti-hypertensive medications for longer may show reduced decline to DTI indices of WM (Bender, Daugherty & Raz, 2012). In addition, use of this variable would also include those participants who were newly diagnosed and treated at follow up, but not at baseline.

Blood pressure. Trained laboratory staff used an auscultatory method to measure blood pressure with diastole phase V for identification of diastolic pressure (Pickering et al., 2005). Trained experimenters took measurements as participants sat in a quiet, climate-controlled room on three separate days, normally one to two weeks apart. The values were averaged across measurements to obtain the mean systolic and diastolic pressure.

The Detroit Medical Center hospital laboratory analyzed blood samples collected by a trained phlebotomist from participants following a 12-hour overnight fast. Assays were performed to measure C-reactive protein (CRP), and homocysteine levels.

Lipid Panel. Hospital laboratory staff employed a direct cholesterol oxidase/cholesterol esterase method to measure triglyceride, high-density lipoprotein (HDL) and total cholesterol; reference ranges for total cholesterol and HDL were 100-199 mg/dl and >39 mg/dl, respectively. Hospital laboratory staff used direct measures to calculate low-density lipoprotein (LDL) cholesterol level (in mg/dl) as follows: LDL = Total cholesterol – HDL – (Triglycerides/5).

Blood Glucose. Laboratory staff measured whole blood glucose levels by the standard

enzymatic glucose oxidase method. Blood glucose levels cut-offs were: 70 mg/dl (3.9 mmol/L) > 126 mg/dl (_7.0 mmol/L).

Waist-to-hip ratio (WHR): At the time of blood sampling a trained research staff member measured participant waist and hip circumference using a fabric measuring tape. These values were used to calculate the ratio between the two.

Cognitive Tests

Speed of Perceptual-Motor Processing (Speed)

An experimenter administered the Letter Comparison and Pattern Comparison tests (Salthouse, 1996) to gauge speed of perceptual processing. The experimenter instructed participants to determine whether pairs side-by-side of letter strings or line patterns were the same or different, as quickly and as accurately as possible. Both tests included two pages of items. Participants completed as many items as possible in 30 s provided per page. The performance index for each test is the total correct for both pages, divided by time for completion (# correct / 60 s). Reliabilities for letter and pattern comparison are estimated to be .77 and .87, respectively (Salthouse & Meinz, 1995).

In addition, the mean response times (RTs) from the 1-back verbal and non-verbal trials provide additional indices of speed of perceptual-motor processing.

Executive function (EF)

EF – Inhibition. Participants completed a paper version of the Stroop task (Stroop, 1935; Salthouse & Meinz, 1995). An experimenter first presented the participant with a sheet containing 20 items organized in two columns; each item was surrounded by a 13.5 × 19 mm rectangle. In each test, experimenter instructed the participant to respond as quickly as

possible, starting with the left column, before continuing with the right column. In the color neutral (CN) subtest participants named the color of the ink of strings of six Xs printed in red, green, yellow, and blue ink. In the color incompatible (CI) subtest, the words 'red,' 'green,' 'blue,' and 'yellow' were presented but in incongruently colored ink; participants were instructed to name the color of the ink, rather than the word. Participants completed two versions of each subtest using alternate forms. Interference scores were calculated as the difference between mean response times for CI and CN subtests. This task has an estimated split-half reliability of .72 (Salthouse & Meinz, 1995).

EF – Working Memory: Size Judgment Span. In the task originally described by Cherry and Park (1993) participants are required to maintain representations in working memory, compare them based on semantic features, and re-order the items in ascending physical size for verbal report. In each trial, the experimenter reads a list of items, and there are three trials per set. First, two items are presented per trial with and the number of presented items increases by one upon successful completion of at least two trials per set. The task is concluded when the participant correctly answers fewer than two trials correctly. The task's estimated reliability coefficient is .79 (Cherry & Park, 1993).

EF – Working Memory: Spatial Recall. Working memory was also assessed with a modified, computerized version of the task described by Salthouse (1974, 1975; Salthouse, Kausler, & Saults, 1988). A series of 5 × 5 matrices were presented on a computer screen, each with seven darkened cells. Participants received response forms that included three pages of two columns of five blank matrices. An experimenter told participants that after seeing each matrix, they were to draw an X in each of the seven cells on corresponding blank

matrix on the paper response forms; the experimenter instructed participants to always make seven X marks, guessing if necessary. Participants first completed five practice trials, and then 25 test trials. The index of performance was the average number correct across the 25 test trials. Cronbach's alpha (α) was .89, as computed across the 25 test trials.

EF – Working Memory: Listening Span. In the Listening Span (LSPAN; Salthouse, Mitchell, Skovronek, & Babcock, 1989) task, participants are required to listen to the experimenter reading simple sentences out loud, answer a multiple choice question about the sentences, and freely recall the final word of each sentence, in order of presentation (see Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998 for a full description). The LSPAN contains seven blocks, each with three trials. Participants start with one sentence per trial in the first block, and the number of sentences increases by one for each successive block. After all items in a trial have been presented, the experimenter instructs participants to write down as many of the final words as possible, maintaining the original presentation order. In addition, for trials to be counted as correct participants must correctly answer the accompanying multiple-choice questions. Participants are awarded one point for each correctly recalled and ordered final word that is accompanied by a correct response. One performance index of the LSPAN is the absolute span (AS), calculated as the total number of correct trials from the where at least two out of three trials were correct, starting with the first block. The AS has been used previously (Raz, et al., 1998) as a measure of working memory capacity.

EF – Working Memory: n-Back Tests. Working memory storage and maintenance was assessed with two computerized *n*-back tests using verbal and non-verbal materials (modeled after Dobbs & Rule, 1989; Hultsch, Hertzog, & Dixon, 1990). Verbal and non-verbal tests

presented single-digit numbers and abstract shapes, respectively, on a 17-inch monitor. The 1-, 2-, and 3-back subtests were separately presented, with subtest order counterbalanced in a Latin square across participants. Following presentation of all items in each trial, participants selected the item presented in the position specified by the given subtest. The performance index for both tasks is the number of correct responses (out of 20). The tasks' estimated reliability coefficients are .91 for the verbal and .88 for the nonverbal tests (Salthouse, Hancock, Meinz, & Hambrick, 1996). Because performance is often at ceiling for 1- and 2-back tasks, performance on the 3-back task provides greater meaningful variance.

EF – Task Switching. We used a computerized test (Salthouse, Fristoe, McGuthry, & Hambrick, 1998) to assess participants' ability to switch between stimuli and between tasks. The program instructed participants to associate specific computer keys with different stimulus properties; the program serially presented participants with stimuli (digits), and participants made the appropriate keyboard response based on the changing stimulus property. In a single switch task, participants switched between indicating left or right stimulus presentation. The dual switch tasks required participants to switch between indicating if a number was more ore less than 5 and whether the digit was odd or even. For both right/left and more/odd tasks, the costs due to switching were calculated as the difference in accuracy (total errors) between switch and non-switch trials.

Episodic Memory (EM) – Free recall: Word list. Experimenters administered a task designed by laboratory staff in Visual Basic that presented lists of 16 nouns, for 3 seconds each, on a computer screen. After all words were presented, participants audibly counted backwards by threes from a random 900 number for one minute in order to prevent rehearsal.

Next, the experimenter asked the participants to name out loud as many words as they could remember, in any order, while the experimenter recorded the responses. The process was repeated with the same list toward the end of the testing session, approximately 90 minutes after the first administration. Participants also completed the noun recall task using a new list of words on a separate testing occasion.

EM – Free Recall: Prose Recall. An experimenter administered the logical memory subtest of the Wechsler Memory Scale (WMS-R; Wechsler, 1987). Participants listened as the experimenter read two short narratives. The experimenter asked participants to freely recall all the information, verbatim, both immediately following each presentation and after a 20-minute delay. This measure has an estimated split-half reliability of .74 (Elwood, 1991)

EM – Recognition: *Picture-Name Associations.* Participants completed the Memory for Names subtest of the Woodcock-Johnson Psychoeducational Battery-Revised (Woodcock & Johnson, 1989). Participants serially viewed novel visual stimuli, cartoons depicting 'space creatures,' and listened as the experimenter stated the creature's name consisting of one- and two-syllable nonsense stimuli. Following each new item presentation, participants viewed a page containing multiple pictures and pointed to each previously studied item after the experimenter stated its name. The experimenter provided the correct answer for incorrect responses during the immediate testing phase. There were 72 possible correct responses. Following a 20-minute delay, the experimenter showed the participant 12 pages, each with 12 space creatures; using its previously learned name, the experimenter asked the participant to point to each space creature. Unlike the immediate testing phase, only a single recognition judgment was requested for each presentation of a given page, and the experimenter provided

no feedback on performance accuracy. The maximum score possible in the delayed testing phase was 36. The total number of correct responses for immediate and delayed cued associative memory tests formed the two performance indices for the task. Both immediate and delayed tests have estimated reliabilities of .91 (29).

EM – Recognition: Word pairs. An experimenter administered a recognition test for associative recognition using word pairs (see Bender, Naveh-Benjamin, & Raz, 2010 for details; Naveh-Benjamin, 2000). The task used an intentional encoding condition. All participants received instructions to study and remember both the individual words and the pairs, and an experimenter informed that both would be tested. The task used testing software designed in-lab using Visual Basic. Each participant viewed 26 pairs of unrelated words, presented at a rate of 5.5 s per pair with a 200 ms inter-stimulus interval. To minimize rehearsal following the study phase, the task presented participants with a randomly generated 900 number instructed them to count backwards by threes for 60 s. Participants then completed separate single item, yes/no recognition tests for items (individual words) and associations (word pairs); test order was counterbalanced across the sample. Both item and associative tests included 16 trials. Item test trials presented 16 individual words (8 targets, 8 foils), and the associative test presented 16 pairs (8 intact pairs, 8 recombined pairs). For each trial, participants indicated via keyboard button press if a word had been presented at study or not or if pairs were intact or recombined. After completing both tests the process was repeated with a second list of 26 new word pairs. The lists for each participant were randomly assigned out of six possible lists. At the second wave of testing, roughly two years, one of the lists was repeated from the initial administration (repeated list) and the other was a new list

containing novel stimuli (non-repeated list).

Fluid reasoning (Gf). Participants completed two tests of fluid reasoning previously used in studies of aging and lifespan development (Raz et al., 1998, 2008; Rabbitt & Lowe 2000; Schretlen et al., 2000), the Cattell Culture Fair Intelligence Test, form 3B (CFIT, 3B; Cattell & Cattell, 1973) and Letter Sets Test (parts 1 and 2) from the Educational Testing Service Factor--Referenced Test Kit (Ekstrom et al. 1976). A test of nonverbal reasoning, the CFIT comprises four subtests; each subtest includes 10 to 14 items of increasing difficulty (see Raz et al., 2008 for a complete description of the task). The total number correct and computed standardized IQ scores for the CFIT (Cattell & Cattell, 1973) are the performance indices for this task. The Letter Sets Test included two pages, each with 15 items. Each item consists of a row of five sets of 4-letter strings; the task instructed participants to identify the rule common to four out of the five sets, and mark the set that does not match the rule. The task provided participants seven minutes to complete each page. For each incorrect response, 0.25 point is deducted from the total number correct to yield the performance index.

Data conditioning

We used Microsoft Excel to calculate age in months by subtracting the date of birth from date of scan and dividing that value by the quotient of 365 (days) / 12 (months). For the word pair tests, we used hit rate and false alarm rate data to compute *A'*, a nonparametric index of discriminability (Pollack & Norman, 1964; Stanislaw & Todorov, 1999; Stewart, 2002). We applied an arcsine transformation to *A'* scores to correct for significant skewness in their distributions. Similarly, we corrected skewness in the distributions of several cognitive variables by applying a log-transformation. These included the LSPAN-AS, letter

and pattern comparison speed, response times (RTs) from the *1*-back verbal and non-verbal tasks, reading speed times from the Stroop task, and switch costs indices from the right/left and more/odd switching tasks. We similarly applied a log-transformation to vascular risk variables homocysteine, CRP, fasting blood glucose, and triglycerides to eliminate skewness in the distributions of those variables. In addition, we multiplied the three diffusivity indices by 1000 to bring those values closer to FA, and to reduce the number of decimal places that the software used for latent modeling must estimate, and thus reduce floating-point errors.

We have found that in Mplus software (Muthén & Muthén, 2012), use of standardized data often results in superior model fit as the software does not appear to handle highly varying scales or variance across multiple variables. Therefore, all data were standardized to z-scores for analysis. Furthermore, in order to use standardized scores to calculate latent difference scores DTI, cognitive, and VR data from the second occasion of measurement were standardized to the first measurement occasion. That is, we calculated z-scores for the second measurement occasion using baseline means and standard deviations.

In select cases, WMH/CSF masking resulted in a relative dearth of coverage for six ROIs. Cases in which the number of voxels was < 3 standard deviations from the mean were excluded from analysis. Thus, one case was removed from analysis of RLIC, CCsplenium, and CBd, two cases were excluded from models of CBv and ILF, and three cases were removed from analyses of uncinate fasciculus (UF) using WMH/CSF-masked data.

Data Analysis

The present study employed a structural equation modeling (SEM) framework to assess two-year change in cognitive abilities, VR factors, DTI indices, and relations thereof.

First, for each of the four DTI indices we fit a series of univariate, two-occasion latent difference score models (LDM; McArdle & Nesselroade, 1994) in Mplus 7 (Muthén & Muthén, 2012). The SEM-based LDM analyses used maximum likelihood (ML) estimation and treated missing data as missing completely at random (MCAR). Although brain volumetric measures appear stable over time in such models (Raz et al., 2005, 2008) this has not been previously established for DTI indices. We fit separate univariate LDMs for each of the 14 ROIs that sampled from left and right hemispheres. We fit separate models for non-masked and WMH/CSF-masked data. LDMs for each region used left and right hemisphere mean values as dual indicators for each occasion (Raz et al., 2005, 2008, 2010). Fornix was not included in the LDM analysis as its presence on the midline of the WM skeleton precluded division into multiple observed variables via hemispheric masking. Initially, to test for a significant mean difference between factors representing each measurement occasion or significant variance therein, we fit LDMs to the data without additional covariates.

We assessed model fit using several indices: the comparative fit index (CFI), and Tucker-Lewis Index (TLI) compare model fit to that of a null model, and values of .95 were cutoffs for both the CFI and TLI. For chi-square (χ^2) tests of model fit, a nonsignificant ($p > .05$), smaller χ^2 value indicates acceptable fit in comparison to a null model. A related, more informative fit statistic, χ^2 divided by degrees of freedom (Jöreskog & Sörbom, 1993) used a fairly conservative cut-off value of ≤ 2.0 (Mueller, 1996). Additional goodness-of-fit indices included root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), measures of model misspecification and explained variance,

respectively; for both the RMSEA and SRMR, acceptable fit was indicated by values of .08 and below.

In each LDM (Figure 1), a factor was specified for each measurement occasion with the factor loading for the right hemisphere fixed to zero. The factor loading for the left hemisphere was freely estimated, but an equality constraint was specified between left hemisphere factor loadings for both occasions of measurement. We fixed the intercepts for all four indicators at zero, although in a few select cases, individual paired indicator intercepts were freed if this improved model fit substantially. Model specification also estimated variance for the four observed indicators, and the auto-correlated residuals for each indicator between measurement occasions. The latent difference score factor was obtained by specifying a loading from the Time 2 factor, fixed to one and a regression path from the T1 factor to the Time 2 factor, also fixed to one. The model specified free estimation of variance / residual variance for the T1 factor and difference score factor, while constraining variance / residual variance for the Time 2 factor to zero. Similarly, model specification freely estimated the means for the Time 1 factor and the difference score factor while constraining the mean for the Time 2 factor to zero. In addition, model specification freely estimated the auto-correlated residuals between the Time 1 and latent difference score factors, but constrained the auto-correlated residuals between the Time 1 and Time 2 factors to zero.

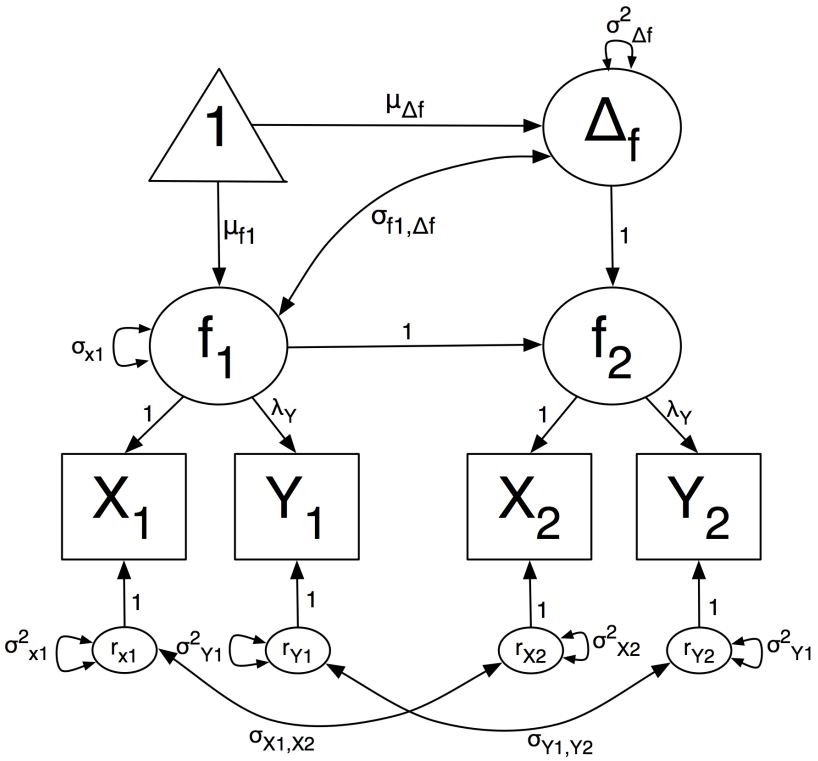


Figure 2. Schematic diagram of a generic univariate LDM measurement model used in analysis of DTI and cognitive data. The squares symbolize observed variables and the larger circles represent latent variables; f1 and f2 represent the latent variable at Time 1 and Time 2, respectively. The triangle signifies means and intercepts for the latent difference factor. The observed indicators X and Y refer to right and left hemisphere measures from the same anatomical region. Auto-correlated residuals for observed variables are kept equivalent between occasions. The factor loading for the X observed variables is fixed at 0, and an equivalency statement is imposed the factor loadings for Y to maintain factorial invariance across occasions. For additional information on two-occasion latent difference models, see McArdle and Nesselroade (1994).

The paramount concern with this modeling approach is the need to establish metric invariance. Metric invariance (Meredith, 1964) refers to a statistical precept necessary for longitudinal modeling of LVs: that the relationship between manifest and latent variables must remain constant between measurements – necessary to confidently interpret differences at the latent level. Metric invariance can be tested in Mplus by simply constraining the factor loadings for the two occasions to be equivalent. If there is metric invariance over time, the model provides a good fit for the data. However, if model fit borders on acceptable, there are small steps that can be taken in an attempt to modify the models. These include freeing factor

covariances, factor means, or residual variances, fixing factor variances to 1, and fixing baseline factor means, individual paired indicator intercepts, and auto-correlated residuals between the baseline and difference score factors to zero. In cases in which it does not reduce model fit, degrees of freedom can also be gained by imposing additional equality constraints on observed indicator intercepts and variance/residual variance, as well as on auto-correlated residuals for a given indicator between measurement occasions. If model fit is still unacceptable and can only be improved by removal of the equality constraints on the Time 1 and Time 2 factor loadings, then there is significant measurement variance over time and change in that cognitive domain cannot be modeled.

Next, a new series of LDMs was performed as before, but with the addition of three covariates: baseline age in months, delay between scans in months, and self-reported years taking hypertension medications at follow up; all three covariates were standardized to z-scores using their own respective means and standard deviations. Univariate LDMs with covariates were specified to include regression paths from all three covariates to the latent difference score factor, and correlations between the Time 1 factor and both age and years on HBP medications, while constraining residualized correlations between the Time 1 factor and delay to zero. We respecified each model to constrain to zero those covariate model parameters that produced non-significant ($p > .4$) parameter estimates close to zero.

Cognitive and VR CFAs & LDMs

We used a similar LDM framework to model latent change in different cognitive domains. However, we first specified confirmatory factor analyses (CFAs) on Time 1 (T1) data in order to identify the factor structure for each construct of interest, and to reduce the

number of manifest indicators required to form latent factors, and change therein. We fit separate CFAs for each domain of cognitive performance, as well as for VR factors.

CFAs

In order to determine the underlying factor structure of cognitive tests for a given domain, we fit separate CFAs by cognitive domain, using performance indices from T1 tests of speed, EF, EM, and Gf. The factor structure then formed the basis of measurement models for subsequent LDM analysis. We removed from subsequent models those indicators that did not load well onto a given factor. The best factor structure was determined by proportion of variance accounted for in the LV, and by evaluating goodness-of-fit indices RMSEA, SRMR, Chi-square, and CFI/TLI.

The scores from the Letter Comparison and Pattern Comparison tests and the 1-back verbal and non-verbal RT scores served as observed indicators for the CFA on speed. A separate EF-working memory CFA included spatial recall score, LSPAN-AS score, and number of correct responses on the 3-back verbal and non-verbal tasks as observed indicators Observed indicators for the CFA on EF included the two switching cost indices and the interference score from the Stroop color word task. Manifest variables for the CFA on Gf included the four CFIT-3B subtest scores and the scores from the two forms of the Letter Sets Test. Observed indicators for the CFA on EM included the number of items recalled from the initial presentation and the non-repeated lists from the word list test, the WMS-R immediate and delayed recall scores, the scores on the WJ-R memory for names immediate and delayed subtests, and the *A'* scores for items and associations from the two lists of the word pair recognition task. The *A'* scores included performance on the lists repeated from baseline (i.e.,

repeated list) and the lists containing novel stimuli at Time 2 (i.e., non-repeated list).

Once we established CFAs with acceptable fit for each domain, we specified LDMs using the matching variables from Time 2 (T2). From the domain-specific cognitive LDMs, we included those models with acceptable fit, metric invariance between measurement occasions, and significant variance in the latent mean difference score in multivariate models of DTI-cognition and VR-DTI-cognition change. For models that did not demonstrate metric invariance, problematic indicators were removed one at a time. For some domains or tasks, a univariate cognitive LDM was fit using a minimum of two indicators per measurement occasion.

A VR CFA using observed indicators systolic and diastolic blood pressure, WHR, and levels of triglycerides, LDL, HDL, CRP, and homocysteine. Unlike the cognitive measures chosen to assess defined cognitive constructs, VR is not inherently a unitary construct. Thus, the CFA on VR indicators may yield a 1-, 2-, or 3-factor structure as the best fit for the data, depending on the correspondence among these factors in this sample. We also fit a CFA for a metabolic syndrome factor using observed indicators systolic BP, triglycerides, WHR, and fasting blood glucose values.

Crossover relations between DTI and Cognition and VR

Next, those univariate DTI models with covariates that showed significant residual variance in the change score factor formed the basis for subsequent multivariate LDM analyses. Our first goal was to test for models in which individual differences in change in DTI-derived WM measures were related to cognitive performance, and where variability in cognitive change predicted individual differences in DTI indices of WM integrity. These

bivariate LDMs included directional (i.e., regression) paths from T1 DTI measures to the cognitive change factor (i.e., the latent difference score for the cognitive factor), from change in DTI (i.e., the latent difference score for the DTI factor) to T2 cognition, from the T1 cognition factor to change in DTI, and from change in cognition to T2 DTI. For DTI indices whose univariate LDMs revealed no significant variance in change, we excluded the two paths including change in DTI from and to T1 cognition and change in cognition, respectively.

False discovery rate (FDR) correction. In light of the large number of models fit to the DTI data, we subjected significance values for latent difference factors and variance of mean difference from the univariate LDMs with and without covariates to correction for false discovery rate (FDR; Pike, 2010).

CHAPTER 3

RESULTS

Univariate LDMs – No Covariates.

A total of 112 univariate LDMs were successfully fit to the data for all 14 ROIs containing bilateral indicators, for all four DTI indices, and for both non-masked and WMH/CSF masked data. We specified all models to converge free of errors, while maintaining factorial invariance. In addition, we calculated Cohen's *d* measures of effect size for the mean difference parameter estimates by dividing the mean latent difference by the square root of variance for baseline DTI factor. The results are shown in Table 2. All 112 models fit well according to multiple goodness of fit indices: $\chi 2$ range = 0.076 to 12.845, all $p \geq .117$, CFI $\geq$ 0.947, TLI $\geq$ 0.960, RMSEA values $\leq$ 0.79. However, in five instances, acceptable model fit was only possible by freely estimating this indicator intercept: models on non-masked data for PLIC FA and UF MD, and models using WMH/CSF masked data for ALIC DR, and SLF DR and MD. Spaghetti plots detailing patterns of change in DTI over time are presented in Appendix A.

Results from the univariate LDMs for DA are presented in Table 2. LDMs fitted to the non-masked data revealed significant mean difference in DA in all regions except for FMaj and CC body. Furthermore, only CC genu and FMin demonstrated significant mean increases in DA, whereas all other regions with significant change exhibited declines. In addition, there was significant variance in change in DA in ALIC, CC body, FMaj, FMin, RLIC, and SLF. LDMs performed on the data masked for WMH/CSF showed a similar pattern; however, following masking and FDR correction, change in CC splenium was no longer significant,

and variance in change was no longer significant in ALIC and FMAJ. In the models for non-masked CC genu and WMH/CSF-masked ILF we respecified models by constraining variance in change to zero to eliminate a psi matrix error resulting from negative parameter estimates close to zero.

Table 2. Univariate model results without covariates - DA

ROI	Mean Δ	*p*	Cohen's *D*	Δ Var.	*p*
Non-masked					
ALIC	-3.995	.000*	-0.315	2.221	.026*+
CBd	-2.633	.008*	-0.374	1.892	.059
CBv	-5.723	.000*	-0.908	0.244	.807
CC body	1.471	.141	0.164	3.839	.000*
CC genu	3.070	.002*	0.415	–	–
CC splen.	-2.196	.028*+	-0.228	0.585	.559
FMaj	0.164	.870	0.023	3.073	.002*+
FMin	2.842	.004*	0.222	3.849	.000*
IFOF	-3.388	.001*	-0.303	0.749	.454
ILF	-4.188	.000*	-0.520	1.005	.315
PLIC	-7.318	.000*	-0.723	1.431	.152
RLIC	-7.530	.000*	-0.809	2.361	.018*
SLF	-5.374	.000*	-0.340	3.796	.000*
UF	-3.387	.001*	-0.337	1.321	.187
WMH/CSF-Masked					
ALIC	-4.775	.000*	-0.337	2.088	.037
CBd	-2.260	.024*	-0.359	1.549	.121
CBv	-6.955	.000*	-1.152	1.767	.077
CC body	0.533	.594	0.058	4.117	.000*
CC genu	3.304	.001*	0.496	1.395	.163
CC splen.	-1.960	.050	-0.217	0.922	.357
FMaj	1.017	.309	0.160	1.910	.056
FMin	2.221	.026*	0.171	3.734	.000*
IFOF	-3.725	.000*	-0.359	1.346	.178
ILF	-4.138	.000*	-0.428	–	–
PLIC	-6.999	.000*	-0.751	0.509	.611
RLIC	-8.341	.000*	-0.890	2.263	.024*
SLF	-5.699	.000*	-0.360	3.469	.001*

UF	-3.676	.000*	-0.377	1.795	.073

Notes: Mean Δ and Δ Var. values are estimates divided by standard error; − = parameter constrained to zero
* = significant following FDR correction
+ = not significant in other masked|non-masked data

Table 3. Univariate model results without covariates - DR

ROI	Mean Δ	*p*	Cohen's *D*	Δ Var.	*p*
Non-masked					
ALIC	0.982	.326	0.057	2.758	.006*
CBd	-2.216	.027*+	-0.238	3.241	.001*
CBv	-3.064	.002*	-0.402	–	–
CC body	-2.278	.023*	-0.119	3.997	.000*
CC genu	5.136	.000*	0.329	–	–
CC splen.	1.734	.083	0.133	2.899	.004*+
FMaj	4.654	.000*	0.284	0.863	.388
FMin	7.194	.000*	0.382	4.569	.000*
IFOF	2.872	.004*	0.141	2.824	.005*
ILF	-1.004	.315	-0.082	1.287	.198
PLIC	-6.792	.000*	-0.737	3.661	.000*
RLIC	-4.323	.000*	-0.326	–	–
SLF	0.125	.901	0.024	2.39	.017*
UF	-1.302	.193	-0.123	2.293	.022*
WMH/CSF-Masked					
ALIC	–	–	0.000	2.882	.004*
CBd	-1.753	.080	-0.158	3.061	.002*
CBv	-2.862	.004*	-0.374	0.527	.598
CC body	-4.707	.000*	-0.297	5.773	.000*
CC genu	4.926	.000*	0.220	0.359	.720
CC splen.	2.350	.019*+	0.153	2.106	.035
FMaj	5.445	.000*	0.335	0.477	.634
FMin	5.517	.000*	0.311	4.655	.000*
IFOF	2.453	.014*	0.130	2.825	.005*
ILF	–	–	–	0.427	.670
PLIC	-7.988	.000*	-0.746	3.581	.000*
RLIC	-4.222	.000*	-0.331	–	–
SLF	-0.138	.891	-0.006	2.487	.013*
UF	-1.478	.139	-0.128	2.219	.026*

Notes: Mean Δ and Δ Var. values are

estimates divided by standard error; − = parameter constrained to zero
* = significant following FDR correction
+ = not significant in other masked|non-masked data

The results from univariate LDMs on DR are presented in Table 3. Analyses of non-masked data showed significant mean change in nine out of 14 regions. There was no significant mean change in DR in ALIC, CC splenium, ILF, SLF, or UF. There was significant variance in change in 10 regions. However, FMaj and ILF showed no significant variance in change. To improve fit or eliminate errors arising from negative variance we constrained the variance in change parameter to zero in DR models for CBv, CC genu, and RLIC. A similar pattern was found in models of WMH/CSF-masked data. Only change in CBd was significant in non-masked but not masked data, whereas change in splenium was significant in masked but not non-masked data. Similarly, variance in change in DR was significant in the LDM for non-masked CC splenium, but not after masking. As with DA, the direction of mean DR change varied by region. We found that mean DR increased in CC genu and splenium, FMaj, FMin, and IFOF, but decreased over time in CBd, CBv, CC body, PLIC, and RLIC. There were psi matrix errors in models for non-masked data in CBv, and CC genu, and both masked and non-masked RLIC models, resulting from negative variance in change parameter estimates. If those estimates were close to zero, we addressed the error by respecifying model syntax to constrain the parameters to zero. Similarly, we addressed psi matrix errors for WMH/CSF-masked data ROIs ALIC and ILF by constraining the mean difference parameter to zero.

Table 4 presents results from univariate models without covariates on non-masked

and WMH/CSF-masked FA data. LDMs for half of the ROIs showed significant two-year mean change in FA, including ALIC, CC body and genu, FMaj, FMin, IFOF, and PLIC. In addition, we found significant variance in change in FA in ALIC, CBd, CC body, FMin, IFOF, PLIC, and UF. The LDMs on the WMH/CSF-masked FA data showed the same pattern, with a few exceptions. Following masking, there was no longer significant mean change in FA in ALIC, and variance in change was also non-significant in ALIC and PLIC.

Table 4. Univariate model results without covariates - FA

ROI	Mean Δ	*p*	Cohen's *D*	Δ Var.	*p*
Non-masked					
ALIC	-3.027	.002*+	-0.212	2.301	.021*+
CBd	0.928	.354	0.112	3.584	.000*
CBv	–	–	0.000	0.100	.920
CC body	3.139	.002*	0.261	4.924	.000*
CC genu	-5.024	.000*	-0.266	1.673	.094
CC splen.	-1.984	.047	-0.144	2.063	.039
FMaj	-4.089	.000*	-0.256	1.252	.210
FMin	-5.059	.000*	-0.315	4.164	.000*
IFOF	-4.080	.000*	-0.274	3.122	.002*
ILF	-1.156	.248	-0.139	0.967	.334
PLIC	2.916	.004*	0.267	3.132	.002*+
RLIC	-0.945	.345	-0.107	–	–
SLF	–	–	–	2.146	.032
UF	-0.301	.763	-0.025	2.446	.014*
WMH/CSF-Masked					
ALIC	0.458	.647	0.041	2.121	.034
CBd	0.405	.685	0.048	3.438	.001*
CBv	-0.404	.686	-0.058	0.098	.922
CC body	3.620	.000*	0.326	4.876	.000*
CC genu	-3.291	.001*	-0.158	0.493	.622
CC splen.	-1.338	.181	-0.115	1.578	.115
FMaj	-4.024	.000*	-0.287	1.084	.278
FMin	-4.214	.000*	-0.262	4.126	.000*
IFOF	-3.568	.000*	-0.262	2.959	.003*
ILF	-1.495	.135	-0.183	1.158	.247

PLIC	3.276	.001*	0.321	2.040	.041
RLIC	-1.420	.155	-0.154	−	−
SLF	-1.265	.206	-0.070	2.166	.030
UF	-0.489	.625	-0.042	2.301	.021*

Notes: Mean Δ and Δ Var. values are estimates divided by standard error; − = parameter constrained to zero
* = significant following FDR correction
+ = not significant in other masked|non-masked data

There was a significant mean increase in FA in CC body and PLIC, and significant mean decrease in ALIC (non-masked only), CC genu, FMaj, FMin, and IFOF. We corrected psi matrix errors in the models for non-masked CBv and SLF by constraining the mean change parameter to zero. We addressed similar errors in non-masked and WMH/CSF-masked models of RLIC by fixing the variance in change parameter to zero.

The results of the univariate LDMs for MD showed widespread mean change (Table 5). In analyses of the non-masked data, there was no significant mean change in ALIC, CC body and splenium, IFOF or UF. We observed significant mean change in ALIC, CC body, and UF only in WMH/CSF-masked data. We found significant variance of change in ALIC, CC body, FMin, and PLIC in non-masked data, and also in SLF in WMH/CSF-masked MD data only. There were significant mean increases in MD in CC genu, FMaj and FMin. In contrast, MD showed mean decreases over two years in ALIC (WMH/CSF-masked only), CBd, CBv, CC body (WMH/CSF-masked only), ILF, RLIC, PLIC, SLF, and UF (WMH/CSF-masked only). We eliminated psi matrix errors in models on non-masked data for CBv, ILF, and RLIC, and in WMH/CSF-masked data in CBd, and ILF by constraining the variance in change parameter to zero. By constraining the difference score to zero, we addressed on error in the LDM for WMH/CSF-masked CC splenium.

Univariate LDMs – Covariate Models

The addition of the covariates age (in months), years taking HBP medications, and delay (in months) between scans to the LDMs allowed us to control and account for these influences on variance in change between measurement occasions. In addition, this permitted comparison of mean change effects from the initial univariate LDMs with cross-sectional age differences in the DTI data. However, if any covariate effect was close to zero, that parameter was constrained to zero to improve model fit. As with the initial univariate models, following final specification, all models converged free of errors, and without violating factorial invariance assumptions.

Table 5. Univariate model results without covariates - MD

ROI	Mean Δ	*p*	Cohen's *D*	Δ Var.	*p*
Non-masked					
ALIC	-1.736	.083	-0.115	3.266	.001*
CBd	-4.786	.000*	-0.530	0.157	.875
CBv	-4.746	.000*	-0.760	–	–
CC body	-0.655	.512	-0.051	4.210	.000*
CC genu	4.452	.000*	0.401	–	–
CC splen.	0.160	.873	0.015	2.013	.044
FMaj	3.064	.002*	0.279	2.116	.034
FMin	5.551	.000*	0.394	3.747	.000*
IFOF	-0.008	.994	0.000	0.398	.690
ILF	-3.756	.000*	-0.298	–	–
PLIC	-7.526	.000*	-1.073	2.491	.013*
RLIC	-8.598	.000*	-0.706	–	–
SLF	-4.858	.000*	-0.292	2.157	.031
UF	-1.163	.245	-0.122	2.088	.037
WMH/CSF-Masked					
ALIC	-3.497	.000*+	-0.213	2.353	.019*
CBd	-4.867	.000*	-0.492	–	–
CBv	-4.858	.000*	-0.902	1.277	.202
CC body	-3.738	.000*+	-0.228	4.989	.000*
CC genu	5.381	.000*	0.388	0.976	.329
CC splen.	–	–	0.000	1.421	.155
FMaj	4.472	.000*	0.341	2.785	.005
FMin	4.934	.000*	0.325	3.705	.000*

IFOF	-1.182	.237	-0.062	1.088	.277
ILF	-2.956	.003*	-0.209	–	–
PLIC	-8.083	.000*	-1.131	2.276	.023*
RLIC	-8.522	.000*	-0.726	1.922	.055
SLF	-5.687	.000*	-0.335	2.401	.016*+
UF	-2.940	.003*+	-0.476	1.695	.090

Notes: Mean Δ and Δ Var. values are estimates divided by standard error; – = parameter constrained to zero
* = significant following FDR correction
+ = not significant in other masked|non-masked data

These models revealed significant covariate effects (Table 6) for age on models of ALIC DA and MD (WMH/CSF-masked only in the latter). Although there was a mean two-year decrease in DA and MD across the sample, this effect was reduced with greater age. In contrast, the mean decrease in MD and DR for PLIC (non-masked only in the latter) was greater with old age. In addition, greater number of self-reported years with diagnosed hypertension was associated with greater two-year increases in both DA (non-masked only) and FA (WMH/CSF-masked only) in CC body. Although there was no mean two-year difference in DA in FMaj, more years with diagnosed hypertension was associated with longitudinal reductions in DA. Last, univariate LDMs with covariates showed that the mean decrease in MD, DR, and DA for PLIC (non-masked only in the latter two) and variance in change in MD in non-masked CC body were greater with longer delay between scans.

Table 6. Significant covariate effects on latent difference score factor

ROI	Index	Age	*p*	HBP Yrs.	*p*	Delay	*p*
Non-masked							
ALIC	DA	2.261	.024*	-1.396	.163	–	–
ALIC	MD	1.99	.047	-0.956	.339	–	–
CC Body	DA	0.632	.527	2.332	.020*	-1.104	.270

CC Body	DR	-0.21	.833	-0.338	.735	-2.173	.030
CC Body	MD	–	–	1.919	.055	-2.742	.006*
FMaj	DA	–	–	-2.684	.007*	–	–
IFOF	MD	-2.082	.037	-0.560	.575	-0.628	.530
PLIC	DA	-1.613	.107	–	–	-2.551	.011*
PLIC	DR	-2.493	.013*	–	–	-2.639	.008*
PLIC	MD	-2.666	.008*	–	–	-2.775	.006*
WMH/CSF-masked							
ALIC	DA	2.416	.016*	–	–	–	–
ALIC	MD	2.295	.022*	-0.894	.371	-0.408	.683
CC Body	DA	0.411	.681	2.123	.034	-0.547	.584
CC Body	FA	–	–	2.224	.026*	–	–
FMaj	DA	–	–	-2.585	.010*	–	–
IFOF	MD	-1.974	.048	-0.928	.353	–	–
ILF	DA	–	–	-2.062	.039	0.311	.756
ILF	FA	–	–	-2.035	.042	0.327	.744
PLIC	DA	-1.447	.148	–	–	-2.148	.032
PLIC	MD	-2.276	.023*	-0.173	.862	-2.301	.021*

*= Significant after FDR correction

In addition, the univariate LDMs with covariates also examined bidirectional associations, or auto-correlated residuals between baseline measures of both age and DTI indices. Greater age was associated with lower DA in FMin (both sets), FMaj (WMH/CSF-masked only), and CBd; a negative association of age with RLIC was no longer significant following FDR correction. In contrast, baseline age was positively associated with T1 measures both DR and MD, and negatively associated with FA in all regions except CBv and UF in non-masked data. In LDMs on WMH/CSF-masked data, greater age was significantly associated with DR and FA in all ROIs, and with MD in all but UF. Thus, in DA and DR, masking yielded correlations with age in CBv and UF not apparent in non-masked data. However, the patterns and direction of change did not mirror those of cross-sectional age effects (Table 7).

CFAs – Cognitive Data

We fit a series of CFAs to the cognitive data grouped by domain. Using baseline cognitive data, we fit separate, domain-specific models for tests of executive function (EF), episodic memory (EM), fluid reasoning (Gf), speed of processing (SPEED), and working memory (WkM). Initially, we specified a single factor solution, with all indicators loading onto a common factor. We addressed any issues with convergence by first changing which measure we specified as the first factor indicator, a factor loading that Mplus automatically fixes at 1. In addition, we first specified the models without auto-correlated residuals among the indicators. Next, we used the same approach to fit a two-factor CFA to the data, and specified different factors based on natural divisions among the indicators based on the tasks from which they were taken.

Table 7. Comparison of cross-sectional correlations with age and mean differences

	DA		DR		FA		MD	
	r-age	Mean Δ	*r*-age	Mean Δ	*r*-age	Mean Δ	*r*-age	Mean Δ
Masked								
ALIC	o	−	+	o	−	o	+	−
CBd	o	o	+	o	−	o	+	−
CBv	o	−	+	−	−	o	+	−
CC body	ne	o	+	−	−	+	+	−
CC genu	ne	+	+	+	−	−	+	+
CC splen.	o	o	+	+	−	o	+	o
FMaj	−	o	+	+	−	−	+	+
FMin	−	+	+	+	−	−	+	+
IFOF	ne	−	+	+	−	−	+	o
ILF	ne	−	+	o	−	o	+	−
PLIC	o	−	+	−	−	+	+	−
RLIC	o	−	+	−	−	o	+	−
SLF	ne	−	+	o	−	o	+	−
UF	o	−	+	o	−	o	ne	−

Non-masked								
ALIC	o	−	+	o	−	−	+	−
CBd	−	−	+	−	−	o	+	−
CBv	o	−	o	−	o	o	o	−
CC body	ne	o	+	−	−	+	+	o
CC genu	ne	+	+	+	−	−	+	+
CC splen.	o	−	+	o	−	o	+	o
FMaj	o	o	+	+	−	−	+	+
FMin	−	+	+	+	−	−	+	+
IFOF	ne	−	+	+	−	−	+	o
ILF	ne	−	+	o	−	o	+	−
PLIC	o	−	+	−	−	+	+	−
RLIC	o	−	+	−	−	o	+	−
SLF	ne	−	+	o	−	o	+	−
UF	o	−	o	o	o	o	o	o

Notes: o=nonsignificant effect; − = negative effect; + = positive effect; ne=not estimated

We specified a CFA for EF using five observed indicators: log-transformed more-odd switching costs, Stroop interference costs, and number correct from the *n*-back nonverbal task, the size judgment span task, and the spatial recall task. We multiplied the indices from the Stroop and switching tasks by -1 to align those scales with the other three variables (i.e., higher values = better performance). We specified separate models to compare fit for 1- and 2-factor solutions. The 2-factor solution specified one factor with Stroop interference, switching costs, and spatial recall performance as indicators; number correct from the size judgment span task and nonverbal 3-back tasks served as indicators in the second factor. Comparison of fit indices (Table 8) for the two EF CFAs shows good, albeit similar goodness-of-fit for both models.

Table 8. Goodness-of-fit for cognitive CFAs

Domain	Model	χ^2	df	*p*	χ^2 / df	CFI	TLI	RMSEA	SRMR
EF	1-Factor	4.260	5	.513	0.852	1.000	1.017	0.000	0.030
EF	2-Factor	2.886	4	.577	0.722	1.000	1.032	0.000	0.024
EM	1-Factor	17.530	14	.229	1.252	0.991	0.982	0.051	0.041
EM	2-Factor	16.481	13	.224	1.268	0.991	0.981	0.053	0.032
Gf	1-Factor	3.453	8	.903	0.432	1.000	1.056	0.000	0.021
Gf	2-Factor	3.453	8	.903	0.432	1.000	1.056	0.000	0.021
Speed	1-Factor	14.234	2	.001	7.117	0.919	0.758	0.252	0.049
Speed	2-Factor	0.246	1	.620	0.246	1.000	1.030	0.000	0.005
WkM	1-Factor	6.983	5	.222	1.397	0.984	0.968	0.064	0.035
WkM	2-Factor	4.836	4	.305	1.209	0.993	0.983	0.047	0.029

CFAs for EM included a total of eight indicators, taken from four different tasks. In addition to *A'* scores from the tests of item and associative recognition in the word pair task (using the lists repeated at follow up testing), remaining indicators included measures of immediate and delayed performance on the WMS-r logical memory task, the WJ-R memory for names task, and the noun list recall task. In addition, in order to produce a model with acceptable fit, we specified bidirectional associations between the two indicators from each task. The 2-factor model specified separate factors for recall using logical memory and noun list indicators, and recognition (word pair performance and memory for names indicators). As with the model for EF, both of the memory CFAs fit the data similarly well.

We specified a CFA for Gf using performance on the four subtasks of the Cattell's Culture Fair test, and the number correct from each of the two Letter Sets tasks as indicators. The 2-factor CFA specified separate factors for CFIT and Letter Sets indicators. As shown in Table 8, both models fit the data equally well.

The CFAs for Speed included the performance indices from the Letter Comparison

and Pattern Comparison tasks, and the response times from the 1-back verbal and nonverbal asks. We multiplied the 1-back response times by -1 to align the performance scaling for all four tasks. The 2-factor model specified one factor using the comparison tasks as indicators, and the response times as indicators for the second factor. Whereas, the 1-factor solution did not fit the data well, the 2-factor solution was an excellent fit.

We also specified a CFA for WkM, the indicators for which included the number correct from the verbal and nonverbal 3-back tasks, the size judgment span task, and the spatial recall task, as well as the AS score from the LSPAN task. The 2-factor CFA specified one factor using the two span tasks as indicators and the second factor was formed using the two 3-back tasks and spatial recall task as indicators. Both models provided a good fit for the data, although the 2-factor model was marginally better, based on slightly higher CFI/TLI values, and lower RMSEA and SRMR indices.

Cognitive LDMs

Based on the results of the CFAs, we fit a series of univariate LDMs to the cognitive data, using the same approach as previously employed for the DTI data. With the exception of Speed, all cognitive LDMs used the 1-factor CFA model as the basis for the LDM, as the 1-factor CFAs fit the data as well as the 2-factor models, but more parsimoniously. We fit separate LDMs for each of the two factors from the Speed CFA. It is important to note that although the CFAs for baseline cognitive performance fit the data well, this did not ensure that such a factor structure would be similarly identified in the T2 data, or that factorial invariance assumptions would be met. Thus, if the assumption of factorial invariance was not met for any univariate cognitive LDMs, we re-specified the model using only one of the

factors from the 2-factor CFAs.

We first used the CFA for EF to fit the LDM for that cognitive domain. Indicators for each of the two occasions included log-transformed more-odd switching costs, Stroop interference costs, and number correct from the *n*-back nonverbal task, the size judgment span task, and the spatial recall task. Model fit was acceptable (χ^2[47] = 59.540, p = .104, CFI/TLI = 0.971/0.972, RMSEA = 0.053, SRMR = 0.090). However, Mplus initially estimated a negative value near zero for the variance in change parameter, resulting in a psi matrix error. The final model was produced after constraining the variance in change parameter to zero.

Although the 1-factor CFA for EM fit the data well using paired indicators from four different EM tasks (WMS-r logical memory, noun list recall, memory for names, word pair recognition), the LDM we specified with measurements from both occasions was a poor fit for the data (χ^2 = 584.186, p < .000, CFI/TLI = 0.605/0.615, RMSEA = 0.198, SRMR = 0.172). The poor model fit resulted from the nature of the LDM specification, which precluded specifying auto-correlated residuals among the indicators for each occasion. Therefore, we specified separate models for each task, using the dual indicator LDM approach employed in the univariate DTI models. The LDM using the two indicators from WJ-r Memory for Names (immediate, delayed) was a good fit for the data (χ^2[6] = 7.483, p = .279, CFI/TLI = 0.996/0.996, RMSEA = 0.051, SRMR = 0.017) and showed significant variance in change (estimate/S.E. = 4.050, p < .000). The LDM using the immediate and delayed performance indices from the WMS-r logical memory subtest as paired indicators was a poor fit for the data (χ^2[6] = 85.534, p < .000, CFI/TLI = 0.769/0.769, RMSEA = 0.372,

SRMR = 0.189), and we were unable to respecify the model to achieve an acceptable fit while maintaining factorial invariance.

The univariate EM model using performance on the two noun list recall tasks as paired indicators was a good fit for the data ($\chi^2[3] = 3.824$, $p = .281$, CFI/TLI = 0.994/0.987, RMSEA = 0.056, SRMR = 0.036), but did not show significant variance in change (estimate/S.E. = 1.045, $p = .296$). Last, the LDM using the *A'* scores from the item and associative recognition subtasks of the word pair task (repeated list) was also a good fit for the data ($\chi^2[8] = 7.136$, $p = .522$, CFI/TLI = 1.000/1.006, RMSEA = 0.000, SRMR = 0.062) and showed significant variance in latent change (estimate/S.E. = 2.106, $p = .035$). Although they were not included in the initial CFA on EM, we also modeled the *A'* data generated in response to the stimulus lists from the word pair task that were not repeated from baseline (i.e., the non-repeated lists). As with the univariate LDM on word pair recognition performance using the repeated lists, the model proved a good fit for the data ($\chi^2[8] = 11.010$, $p = .201$, CFI/TLI = 0.944/0.958, RMSEA = 0.064, SRMR = 0.084), but did not demonstrate significant variance in latent change (estimate/S.E. = 0.877, $p = .380$).

We fit an LDM for Gf using the results of the CFA on that construct with four indicators from the CFIT tasks and two from the Letter Sets task. Model fit was only marginally acceptable ($\chi^2[75] = 104.089$, $p = .015$, CFI/TLI = 0.941/0.948, RMSEA = 0.064, SRMR = 0.107). Therefore, we fit separate LDMs for the two tasks from which we took the indicators for the Gf LDM. A univariate LDM for CFIT 3b that included number correct on each of the four subtasks as observed variables was a good fit for the data ($\chi^2[34] = 28.287$, $p = .743$, CFI/TLI = 1.000/1.026, RMSEA = 0.000, SRMR = 0.067), but did not demonstrate

significant variance in latent change (estimate/S.E. = 0.088, p = .930). In contrast, the LDM using number correct from the two Letter Sets tasks as indicators was a good fit for the data (χ^2[6] = 3.664, p = .722, CFI/TLI = 1.000/1.010, RMSEA = 0.000, SRMR = 0.023), and demonstrated significant variance in latent change (estimate/S.E. = 2.113, p = .035).

Because the CFA for Speed revealed the 2-factor solution as superior to a single factor, we first fit a univariate LDM using the Letter Comparison and Pattern Comparison performance indices as observed indicators. Although the final model was a good fit for the data (χ^2[6] = 10.410, p = .108, CFI/TLI = 0.979/0.979, RMSEA = 0.088, SRMR = 0.059), this was only possible by constraining the variance in change parameter to zero to eliminate a psi matrix error that arose due to a near-zero negative estimate in that parameter. Similarly, the LDM for Speed using the response times from the two 1-back tasks as indicators was also a good fit (χ^2[7] = 3.478, p = .838, CFI/TLI = 1.000/1.011, RMSEA = 0.000, SRMR = 0.044), but that model also necessitated constraining the variance in change parameter to zero to mitigate a psi matrix error.

We successfully fit the univariate difference score model for WkM using the same five indicators as used in the CFA; that model provided a good fit for the data (χ^2[48] = 58.300, p = .147, CFI/TLI = 0.980/0.981, RMSEA = 0.047, SRMR = 0.086). However, the variance in latent change parameter was not significant (estimate/S.E. = 1.274, p = .203).

Vascular Risk CFA & LDM

We fit a CFA to the T1 VR indicators of metabolic syndrome: log-transformed triglyceride level, HDL cholesterol level, systolic blood pressure, waist-to-hip ratio, and fasting blood glucose level. HDL was multiplied by -1 to align its scaling with the other

indicators. The CFA was an acceptable fit for the data only if bidirectional paths were specified between HDL and both triglycerides and waist-to-hip ratio ($\chi^2[2] = 0.414$, $p = .813$, CFI/TLI = 1.000/1.102, RMSEA = 0.000, SRMR = 0.011). Moreover, the univariate LDM for metabolic syndrome using the same indicators as those in the CFA and their longitudinal counterparts was also an acceptable fit for the data ($\chi^2[47] = 56.773$, $p = .156$, CFI/TLI = 0.977/0.978, RMSEA = 0.047, SRMR = 0.078), even without the correlational paths specified for HDL. Furthermore, there was significant variance in the latent difference score (estimate/S.E. = 2.125, $p = .034$). However, when we subsequently included the metabolic syndrome LDM in multivariate models to evaluate the effects of change in VR on DTI indices, the models were a poor fit for the data. We found that model fit was only improved by again specifying correlational paths between HDL and triglycerides, which violated basic requirements of the LDM framework. Thus, we specified an alternative LDM for VR without HDL in the model. The new metabolic syndrome LDM without HDL was a good fit for the data ($\chi^2[31] = 21.904$, $p = .886$, CFI/TLI = 1.000/,1.031 RMSEA = 0.000, SRMR = 0.062), but the estimated variance in the latent difference score was no longer significant without HDL in the model (estimate/S.E. = 1.787, $p = .074$).

Multivariate LDMs – DTI-cognition

Three models from the cognitive LDMs demonstrated both acceptable fit and significant variance in the latent difference score. These included the Gf model with two indicators from the Letter Sets task (Letter Sets), the EM model with two indicators from the WJ-R Memory for Names task (Names), and the EM model with two indicators from repeated lists on the word pair task (word pair). Next, for each of these models we specified a

series of multivariate LDMs that combined the cognitive LDM syntax and the syntax from each of the univariate LDMs on the DTI data that showed significant variance in change. These multivariate DTI-cognition models only used the WMH/CSF-masked DTI data. In addition to specifying the previously mentioned covariates of age, years of antihypertensive medication, and delay between scans, we also specified four directional paths between the DTI and cognitive factors in the model. These included a path from T1 DTI to cognitive change score, a path from T1 cognitive factor to the DTI change score, a path from the DTI change score to the T2 cognitive factor, and a path from the cognitive change score factor to the T2 DTI factor. In other words, we modeled the influence of baseline cognition on two-year change in WM and vice versa, and the influence of baseline DTI on cognitive change, as well as the effect of change in DTI and cognition on follow up measures of cognitive performance and DTI, respectively. In addition to these models, we also created multivariate models using those univariate DTI LDMs that did not show significant variance in change, but restricted specifying DTI-cognition relations to not include DTI change. Thus, those models included paths from baseline DTI to change in cognition and from cognitive change scores to T2 DTI.

GF Letter Sets

Higher baseline DA in CC splenium and all three internal capsule ROIs predicted longitudinal declines in Gf (Table 10). Similarly, lower baseline MD in CC splenium predicted greater two-year improvement on the Letter Sets task. In contrast, greater two-year increase in GF Letter Sets predicted lower DR in CC genu at T2. Individual differences in Gf at baseline also explained variance in DTI change scores: higher Gf at baseline predicted

steeper two-year declines in DA in IFOF, DR in ALIC, and MD in ALIC, CC body, and PLIC.

In addition, cross-sectional associations between baseline DTI and cognitive data (Table 9) showed lower T1 Gf was related to higher baseline DR in multiple regions including dorsal and ventral CB, CC genu, FMaj, ILF, and SLF (Table 9). Similarly, greater T1 Gf was associated with higher FA in dorsal and ventral CB, forceps major and minor, and SLF. Higher MD in dorsal CB, CC genu, forceps major, ILF, and RLIC were associated with lower baseline GF. There were no significant associations between baseline Gf and DA in any ROI.

Table 9. Significant baseline associations: DTI-Gf Letter Sets

ROI	Estimate / S.E.	*p*
DR		
CBd	-3.116	.002
CBv	-2.390	.017
CC Genu	-2.140	.032
CC Body	-2.164	.030
FMAJ	-2.456	.014
ILF	-2.347	.019
SLF	-2.059	.040
FA		
CBd	2.807	.005
CBv	2.186	.029
CC Body	2.296	.022
FMAJ	2.265	.023
FMIN	2.217	.027
SLF	2.030	.042
MD		
CBd	-2.268	.023
CC Genu	-1.961	.050
FMAJ	-1.977	.048
FMAJ	-2.101	.036

ILF	-2.124	.034
RLIC	-2.023	.043

Table 10. Significant Paths in DTI-Gf Letter Sets Multivariate LDMs

ROI	T1 DTI → ΔGf	*p*	T1 Gf → ΔDTI	*p*	ΔDTI → T2 Gf	*p*	ΔGf → T2 DTI	*p*
DA								
ALIC	-2.380	.017*	–	–	–	–	0.496	.620
CC Spl.	-2.612	.009*	–	–	–	–	-1.103	.270
IFOF	–	–	-2.552	.011*	–	–	–	–
PLIC	-2.026	.043*	–	–	–	–	1.113	.266
RLIC	-2.255	.024*	-1.663	.096	–	–	–	–
DR								
ALIC	–	–	-2.061	.039*	–	–	–	–
CC Genu	-0.220	0.826	–	–	–	–	-2.148	.032*
MD								
ALIC	–	–	-2.679	.007*	–	–	–	–
CC Body	–	–	-2.225	.026*	0.719	.472	–	–
CC Spl.	-2.216	.027*	–	–	–	–	–	–
PLIC	-1.252	0.211	-2.432	.015*	–	–	–	–

Notes: Values are parameter estimates divided by standard errors; T1=baseline; T2=follow up; ΔGf=latent difference score for Gf; ΔDTI=latent difference score for DTI indices.

Memory for Names

Higher baseline DA in ventral CB predicted greater positive change on memory for names (Names; Table 12). In contrast, higher baseline DA in CC body was associated with smaller rate of improvement in Names. Moreover, lower baseline DR in ventral and dorsal CB and in CC splenium predicted greater longitudinal increase in Names performance. Similarly, lower baseline FA in the ALIC and higher baseline FA in CB ventral and dorsal and forceps minor predicted greater longitudinal improvement on Names. Individual differences in baseline Names performance also explained significant variability in change in FA in the PLIC: higher baseline memory predicted smaller two-year increase in PLIC FA.

In addition, greater two-year decreases in DA in ALIC, and DA and MD CC body predicted better Names performance at follow up. In contrast, greater two-year increases in DR, and smaller increases in FA in CC body predicted better Names performance at follow up. Moreover, greater two-year increases in Names performance predicted higher MD at T2 in CC body and higher DR in SLF, but lower DA in IFOF and lower FA in CBv and CC genu.

Table 11. Significant baseline Associations: DTI-Names

ROI	Estimate / S.E.	*p*
DA		
ALIC	-2.150	.032
FMIN	2.047	.041
RLIC	-2.537	.011
SLF	-2.736	.006
DR		
ALIC	-2.732	.006
CBd	-2.605	.009
CC Genu	-2.116	.034
CC Body	-2.631	.009
FMAJ	-2.823	.005
FMIN	-2.974	.003
IFOF	-3.017	.003
ILF	-2.798	.005
PLIC	-3.499	.000
RLIC	-3.336	.001
SLF	-2.723	.006
FA		
ALIC	2.291	.022
CBd	2.477	.013
CC Body	2.485	.013
FMAJ	2.931	.003
FMIN	3.705	.000
ILF	1.971	.049
PLIC	3.150	.002
RLIC	2.069	.039
SLF	1.952	.051

MD		
ALIC	-3.162	.002
CBd	-2.478	.013
CC Genu	-2.138	.032
CC Splenium	-2.135	.033
CC Body	-2.265	.024
IFOF	-3.123	.002
ILF	-2.540	.011
PLIC	-2.079	.038
RLIC	-3.851	.000
SLF	-3.198	.001

Table 12. Significant Paths in DTI-EM (Memory for Names) Multivariate LDMs

ROI	T1 DTI → ΔEM	*p*	T1 EM → ΔDTI	*p*	ΔDTI → T2 EM	*p*	ΔEM → T2 DTI	*p*
DA								
ALIC	0.922	.357	–	–	-2.078	.038*	0.987	.324
CBv	2.361	.018*	–	–	–	–	-1.336	.182
CC Body	-2.885	.004*	–	–	-2.290	.022*	1.817	.069
IFOF	0.967	.334	–	–	–	–	-2.805	.005*
DR								
CBd	-1.959	.050	1.805	.071	–	–	1.414	.157
CBv	-2.127	.033*	–	–	–	–	0.987	.324
CC Body	-0.951	.342	–	–	2.920	.003*	–	–
CC Spl.	-2.172	.030*	–	–	-1.285	.199	1.123	.261
SLF	-1.457	.145	1.478	.139	-1.569	.117	2.408	.016*
FA								
ALIC	-2.361	.018*	–	–	–	–	–	–
CBd	2.134	.033*	-1.293	.196	–	–	–	–
CBv	3.387	.001*	–	–	–	–	-2.259	.024*
CC Body	–	–	-0.906	.365	-2.848	.004*	–	–
CC Genu	0.333	.739	–	–	–	–	-2.668	.008*
FMin	2.093	.036*	–	–	–	–	-1.349	.177
PLIC	–	–	-2.755	.006*	–	–	–	–
MD								
CC Body	-2.669	.008*	–	–	-2.031	.042*	2.492	.013*

Notes: Values are parameter estimates divided by standard errors; T1=baseline; T2=follow up; ΔGf=latent difference score for Gf; ΔDTI=latent difference score for DTI indices.

Baseline Names performance was significantly associated with T1 DTI measures in numerous regions (Table 11). Higher DA in forceps minor was associated with better baseline Names performance. However, increased DA in ALIC, RLIC, SLF were all associated with reduced baseline Names performance. Similarly, higher baseline DR in CBd, ALIC, PLIC, RLIC, CC genu, CC body, forceps major and minor, IFOF, ILF, SLF were all associated with lower baseline memory for names. In contrast, higher FA in those same regions except for IFOF was related to superior Names performance. Higher baseline MD in all three internal capsule ROIs, all three CC ROIs, dorsal CB, IFOF, ILF, and SLF was associated with worse performance.

Word pair task

The multivariate LDMs for the word pair task revealed numerous instances in which baseline DTI measures predicted two-year change in recognition memory performance (Table 13). Higher baseline DA in CC body, DR in dorsal CB, CC genu and splenium, forceps major and minor, and IFOF, ILF, and SLF, and higher MD in all three CC ROIs and forceps minor were associated with two-year reductions in recognition memory. Similarly, lower baseline FA in dorsal CB, CC genu and splenium, forceps major and minor and PLIC were associated with reduced longitudinal performance on the word pair task. In addition, two-year increases in DA in ALIC and CC body, and MD in CC body were associated with reduced memory performance at follow up. There were no significant paths from baseline word pair performance to change in DTI or from change in word pair performance to T2 DTI.

There were significant bidirectional associations between baseline measures of DTI and word pair performance (Table 14). Higher T1 DA in forceps major and minor was

associated with better baseline memory. Similarly, higher baseline FA in dorsal CB, CC body, forceps minor, and PLIC was associated with superior recognition at T1. Elevated baseline DR and MD in dorsal CB and RLIC and higher PLIC MD were all associated with poorer concurrent recognition performance.

Table 13. Significant Paths in DTI-EM (Word pair task) Multivariate LDMs

ROI	T1 DTI → ΔEM	*p*	T1 EM → ΔDTI	*p*	ΔDTI → T2 EM	*p*	ΔEM → T2 DTI	*p*
DA								
ALIC	–	–	-0.474	.635	-2.479	.013*	–	–
CC Body	-3.069	.002*	-1.296	.195	-3.572	.000*	–	–
DR								
CBd	-2.482	.013*	1.469	.142	-1.468	.142	1.539	.124
CC Genu	-2.411	.016*	–	–	–	–	–	–
CC Spl.	-2.452	.014*	-1.752	.080	–	–	–	–
FMAJ	-2.394	.017*	–	–	–	–	–	–
FMin	-3.135	.002*	–	–	–	–	–	–
IFOF	-1.999	.046*	–	–	–	–	–	–
ILF	-2.246	.025*	–	–	–	–	-0.601	.548
SLF	-1.948	.051	–	–	-1.649	.099	–	–
FA								
CBd	2.530	.011*	-1.275	.202	1.481	.139	-1.415	.157
CC Genu	2.063	.039*	–	–	–	–	–	–
CC Spl.	2.110	.035*	–	–	–	–	-0.742	.458
FMAJ	1.967	.049*	–	–	–	–	–	–
FMin	2.878	.004*	-0.907	.364	0.757	.449	–	–
PLIC	2.532	.011*	-0.383	.702	0.867	.386	-1.172	.241
MD								
CC Body	-2.511	.012*	-1.011	.312	-1.967	.049*	–	–
CC Genu	-2.217	.027*	–	–	–	–	-1.472	.141
CC Spl.	-2.780	.005*	–	–	–	–	0.313	.755
FMin	-2.336	.019*	–	–	–	–	–	–
SLF	-2.011	.044	–	–	-1.482	.138	–	–

Notes: Values are parameter estimates divided by standard errors; T1=baseline; T2=follow up; ΔGf=latent difference score for Gf; ΔDTI=latent difference score for DTI indices;

Multivariate LDMs – DTI-VR

We combined the syntax for the univariate DTI models of WMH/CSF-masked data that showed significant variance in change with the univariate model of metabolic syndrome without HDL cholesterol. Each model estimated a path from baseline VR to the DTI latent difference score, and a bidirectional pathway between baseline VR and DTI factors. Although all models had acceptable fit (for all, $p > .15$, CFI/TLI > 0.970/0.970, RMSEA < .04), none demonstrated a significant path from the T1 VR factor to change in DTI. However, several did show significant cross-sectional relationships between baseline factors for VR and DTI. Higher VR was associated with lower DA in forceps minor and higher DA in RLIC. In addition, higher levels of VR were also associated with greater DR in several regions including dorsal CB, CC body and splenium, forceps minor, IFOF, and SLF, and with greater MD only in CC body. Increased VR was negatively associated with FA in similar regions as DR: dorsal CB, CC body, forceps minor, IFOF, and SLF.

Table 14. Significant baseline Associations: DTI-Word pair

ROI	Estimate / S.E.	*p*
DA		
FMAJ	1.959	.050
FMIN	2.755	.006
DR		
CBd	-2.621	.009
PLIC	-2.357	.018
RLIC	-1.965	.049
FA		
CBd	2.490	.013
CC Body	2.277	.023
FMIN	2.201	.028
PLIC	2.214	.027
MD		
CBd	-2.126	.033

RLIC	-2.219	.026

CHAPTER 4

DISCUSSION

Using the SEM-based LDM framework, we observed significant, reliable change over two years in all four DTI indices. Although the two-occasion LDM approach has been used successfully on brain volumetric data (Raz et al., 2005, 2008, 2010), these results show its utility with DTI data, and over a shorter longitudinal delay. Moreover, whereas cross-sectional associations between age and DTI indices of WM were consistent across brain regions, the direction of mean change varied by DTI index and ROI. Furthermore, numerous regions also demonstrated significant variance in change that permitted multivariate analysis comparing influences of WM change on cognitive performance, and vice versa. The present findings also demonstrate the utility of an atlas-based ROI approach for longitudinal assessment, and the importance of analyzing data without the influence of WMH/CSF.

Comparison of cross-sectional and longitudinal effects in the univariate LDMs (Table 7) supports the suggestion that not only do cross-sectional associations between age, cognitive performance, and their putative biological substrates provide a poor estimate of differences or change over time, but they cannot even constrain the search space for such effects (Lindenberger et al., 2011). For example, whereas age was negatively associated with individual differences in baseline DA in only three regions, change measures told a different story altogether. Although most regions showed mean decline, DA increased in anterior regions CC genu and forceps minor. Similarly, although DR was positively associated with age in almost all regions, only five regions showed a mean two-year increase in DR, while four regions showed a mean decrease. In addition, whereas FA was negatively correlated

with age in all regions using masked data, FA showed a mean increase over time in PLIC and CC body; in contrast, there was a mean decrease in FA only in CC genu, FMaj, FMin, and IFOF. Finally whereas baseline MD was positively associated with age in almost all regions, MD actually showed a mean decline in nine out of 14 ROIs, and only showed mean two-year increase in CC genu, and forceps major and minor. Clearly, cross-sectional associations between DTI measures of WM and age are uninformative regarding actual change.

Moreover, covariate effects showed differential influence of age on WM change. For example, although we observed longitudinal declines in DA and MD in ALIC and PLIC, greater age attenuated declines in the former region but exacerbated declines in the latter. Similarly, DA in FMaj and CC body both showed no significant mean change, but did demonstrate significant variance in change. Greater number of years of treated hypertension was related to greater increase in DA and FA in CC body and attenuated DA increases in FMaj. Thus, both age and VR account differentially account for variance in change DTI indices, across regions.

Effects of masking

We performed the univariate LDMs using both whole brain data, and on WM images with WMH/CSF removed. In some regions, masking appears to have reduced variance in change. In addition, age was correlated with DR and FA in both ventral CB and UF only in models of WMH/CSF-masked data. DA in FMaj was also only significantly correlated with age in masked data. One possibility is smaller regions may contain more noise or less reliable signal than larger WM tracts. In contrast, DA in dorsal CB was only negatively related to age in non-masked data. Therefore, it is unclear whether and to what extent the presence of

WMH may have contributed to the present findings. Alternatively, it is possible that our DTI sampling methodology included the undue presence of fibers from other regions such as CC body, which contained age-related variance not necessarily present in cingulum.

DTI-Cognition

The discrepancy between baseline cross-sectional associations of DTI measures and cognition and longitudinal effects is an important finding of the present study. Whereas cross-sectional associations demonstrated uniformity within each DTI index, the different influences of baseline DTI and cognitive measures on change in cognition DTI, respectively, show no such consistency. It is important to keep in mind that the multivariate models of interactive influence between DTI measures and cognition were based on cognitive models showing significant variance in change. Moreover, the two models of episodic memory both showed mean increases in performance. Therefore, these models are not evaluating the effects of change in DTI on age-related decline in cognitive performance, but rather individual differences in both change and the benefit of repeated testing, or retest effects, as indicated by the variability in the slope of the latent difference score.

The models of latent change in Gf showed significant variance in change, but not mean change over two years. Furthermore, greater two-year increase in Gf was associated with lower DR in CC genu at follow up. More generally, the influence of baseline DA on change in Gf underscored the influence of diffusivity in the three internal capsule ROIs on variability cognitive change. Notably, despite the apparent influence of baseline DA on two-year change in Gf, there were no cross-sectional associations between Gf and DA in any ROI. Thus, the baseline individual differences in reasoning appear to be qualitatively

different from individual differences in change in that ability.

In addition, both of the LDMs for episodic memory (Memory for Names and word pair recognition) demonstrated a negative influence of DA in CC body on two-year change in performance. Higher DA was associated with decreased longitudinal improvement. In addition, models for both memory tasks revealed the negative influence of change in DA in CC body and ALIC on memory performance at follow up, such that greater increase in DA predicted poorer memory at follow up. However, only Memory for Names showed relationships with DA, DR and FA in ventral cingulum, in which higher baseline FA and DA and lower baseline DR predicted longitudinal memory improvements. The models for the repeated stimulus list for the word pair task showed baseline involvement of a similar set of structures, for both DR and FA. These common regions included dorsal CB, CC genu and splenium, and forceps major and minor. Moreover, in the word pair LDMs greater baseline FA and lower baseline DR predicted larger two-year improvement in memory.

In light of both the discrepancies between cross-sectional and longitudinal effects, and the contradictory nature of many of the longitudinal DTI-cognition relationships, unequivocal interpretation of the present findings poses a particular challenge. DTI findings from animal studies suggest DR serves as a proxy for myelin and DA as a measure of axonal integrity (Song et al., 2002, 2005; Sun et al., 2008). However, many of the present findings are inconsistent with this account. If DR is an inverse measure of myelination, why would two-year increases in DR or MD be associated with better performance? Similarly, why would reduced FA, a putative marker of fiber tract coherence, predict better memory performance? One possible answer is that individual differences in cross-sectional DTI

measures may account for different variance in cognition than individual differences in two-year change in DTI. That is, greater age may indeed be associated with reduced FA, but over a two-year delay, smaller changes in FA appear to be wholly independent of age. Similarly, it is possible that intra-individual change in the DTI signal reflects different underlying biological processes than individual differences in the stable signal.

More recently, some have begun to suggest greater caution in the interpretation of DTI data as indicative of underlying tissue properties (Jones et al., 2013; Wheeler-Kingshott & Cercignani, 2009) noting various possible influences on DTI indices beyond simply myelination. Jones et al. (2013) suggest that developmental differences in WM such as increased axonal diameter or reduced packing density could both result in reduced anisotropy. Moreover, estimates suggest that 30-90% of WM voxels contain multiple fibers of different orientation (Behrens et al., 2007; Jeurissen et al., 2013; see Jones et al., 2013 for a review). Therefore, one explanation suggests that the increase in axonal diameter or myelination of a secondary fiber population of different orientations than the principal eigenvector in a given ROI would result in increased DR and reduced FA. This seems more plausible than reductions in myelination driving increases in cognitive performance.

This is not to suggest that over time, some change on the mean level may not occur to reflect cross-sectional age variance, but that such processes are poorly understood, and such differences may plausibly reflect cohort effects. Longitudinal studies must account for these discrepancies, but the timescale needed imposes a particular burden as methods and hardware advance over time and are replaced. Nonetheless, prospective and longitudinal designs should be a new standard for DTI studies of cognitive aging, in order to more clearly understand the

relationship between age-related change in WM and cognition.

Study Limitations

The single-tensor solution may not provide the best characterization of cerebral white matter, particularly in regions in which many voxels include crossing fibers (Jones & Cercignani, 2010; Mukherjee, 2008; Vos, Jones, Jeurissen, Viergever, & Leemans, 2011). Rather, techniques that utilize multiple b-values, or multi-pool methods that evaluate multiple rates of diffusion may provide a better method for disentangling extra-axonal, intra-axonal, and intra-myelin diffusion. At best, DTI provides an imaging modality for quantification of important aspects of underlying white matter microstructure. At worst, DTI is an unreliable correlate of white matter anatomy, and poor interpretations can lead to drawing wrong conclusions and only add noise to the existing understanding of the effects of age on WM and the role of WM change in age-related cognitive decline.

In addition, several factors in the present study may limit the reproducibility and generalizability of the findings. Similarly, although the data appear generally stable, there was no phantom used as a control for the signal as used by Teipel et al. (2010), which would have been preferable. However, overlap within the group between occasions may have mitigated the potentially confounding effects of scanner drift or software or hardware updates or upgrades. That is, participants at the end of the first wave of measurements were being scanned as the earliest participants were beginning to undergo the second measurement. Moreover, measurements made on the same scanner over time appear to be robust to longitudinal drift, even following software upgrades (Takao et al., 2012).

In addition, although sampling from native space data may be preferable in

minimizing interpolation, the differences in standard and native space could create additional noise. That is, skeletonisation is done in 1 mm^3 standard space, and the native DTI data are not. Therefore, the deprojected skeleton is inherently larger in native than in standard space due to differences in voxel size (e.g., in mm 2 × 2 × 3 mm vs. 1 mm^3). However, this largely imposes issues only at the periphery of the skeleton before, which we addressed by applying a slightly more conservative threshold for the skeleton in the final stage of TBSS processing. Moreover, deprojection to native space permitted better evaluation of the correspondence of the deprojected skeleton and skeletonized atlas ROIs with the underlying anatomy in the b_0 and FA images.

Tractography might possibly provide more anatomically specific and less noisy DTI measures than our ROI approach. Despite the care that went into the DTI processing, the use of a probabilistic atlas-based system for anatomical designation is not as specific as tractography. Although ROIs did not overlap, neighboring regions such as ILF and IFOF although containing different variance, may not have the anatomical validity as streamlines. However, as with many archival DTI data, tractography is not advisable with fewer than 32 encoding gradient directions (Jones & Cercignani, 2010). In addition, we were unable to model change of non-lateralized structures like the fornix using the LDM framework as multiple indicators were not available, and all differences could only be evaluated on the manifest level. Moreover, our desire to use CSF/WMH-masked values for analyses made fornix particularly small. Future studies employing more current, optimized imaging standards might be able to better delineate such structures for more reliable longitudinal comparison.

In the present study we have taken great efforts to minimize noise in the data as much as possible. This was done through use of a highly controlled sample, in our DTI processing and sampling procedures, and in our statistical modeling approach. However, it is possible that such a conservative methodology may bias our findings away from some potential discovery of less robust effects that do not meet the same thresholds for latent difference modeling, bootstrapping, and FDR correction.

In addition, we did not have uniform coverage of all age ranges in the sample. Although well screened, the sample was more heavily middle-aged and less representative of younger adults. Also, we did not analyze the sample by different genetic polymorphisms that might explain variability in change in the present sample. That is, even when mean differences were manifest, small effect sizes may reflect an admixture of genetically conveyed influences. For example, it is possible that BDNF val/val carriers may show greater increases in diffusivity over time, in comparison to carriers of the met allele (Chiang et al., 2011). In the present sample, it is possible that modest or null effects may reflect such heterogeneity. Similarly, we did not evaluate differences due to sex, and the effects of post-menopausal hormone changes on DTI measures of cerebral WM have yet to be clearly elucidated. Similarly, the LDM framework precluded specific evaluation of lateralized change effects and their role in explaining variance in cognitive performance and change therein. That is, we did not examine how differential associations between left or right hemispheres may have shown greater associations with memory performance, or two-year improvements in memory.

Future directions

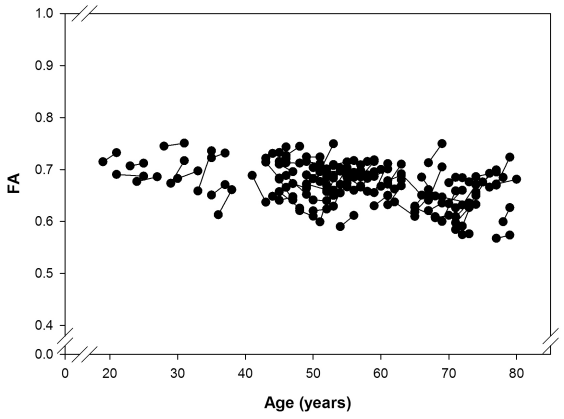
Although we have already identified several ways in which the research could be improved, a whole brain, longitudinal tractography study using a similar sampling approach is a clear next step. Similarly, given the open question of interpretation of change in the DTI signal, a comparison with longitudinal using other WM measures like magnetization transfer imaging, R2*, relaxometry, and myelin water fraction imaging would also provide additional point for comparison that is more anatomically specific. Whereas we evaluated DTI measures with non-masked and WMH/CSF-masked data, a better approach would be to examine change in DTI indices within WMH (Vernooij et al., and model that as a multivariate change factor with change in normal appearing white matter. This would provide a more optimal method for understanding the influence of WMH on

In addition, even though we have used the best statistical framework for assessing change between two occasions, this is inadequate to accurately describe development. Thus, three or ideally four waves or more would be needed to model nonlinear dynamics between age, WM and cognition. Moreover, inclusion of a larger sample of younger participants in the third and fourth decades and over 80 years of age would provide better representation of variance in those life periods as well.

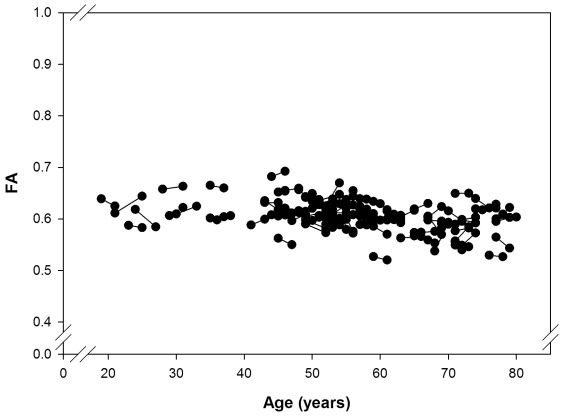
APPENDIX A – Spaghetti Plots of change in DTI indices across regions

Fractional anisotropy (FA)

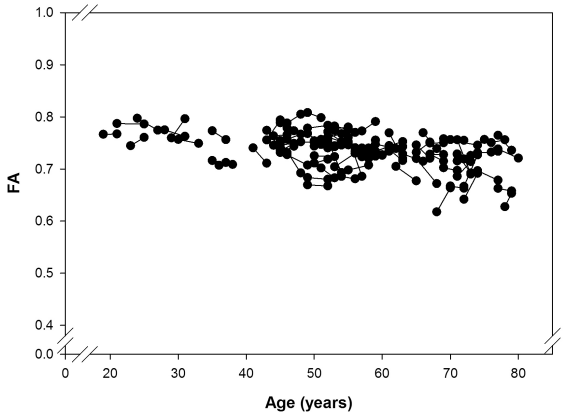
FA Change in Corpus Callosum Body with Age



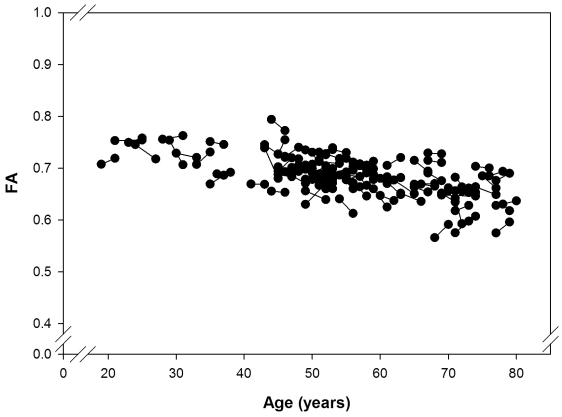
FA Change in Anterior Limb of Internal Capsule with Age



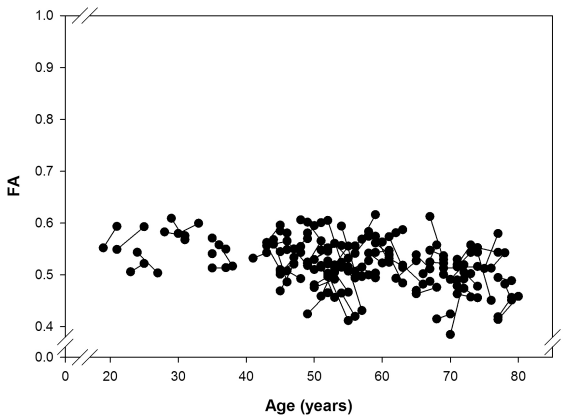
FA Change in Splenium of Corpus Callosum with Age



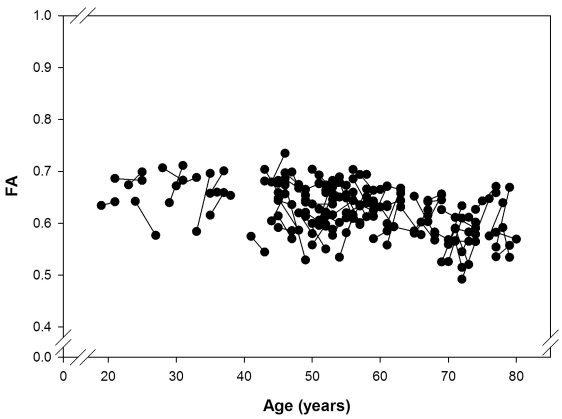
FA Change in Genu of the Corpus Callosum with Age



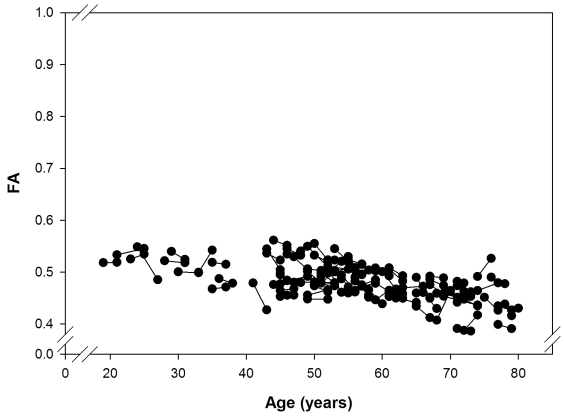
FA Change in Ventral Cingulum Bundle with Age



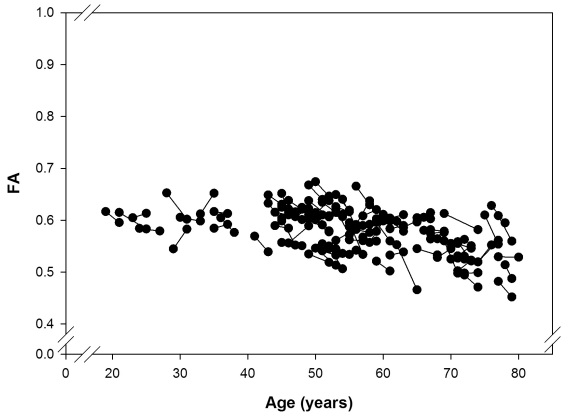
FA Change in Dorsal Cingulum Bundle with Age



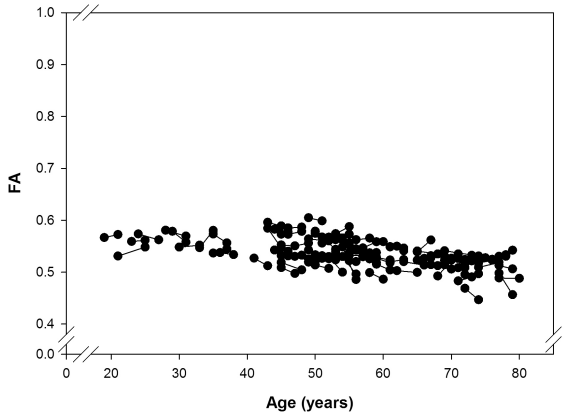
FA Change in Forceps Minor with Age



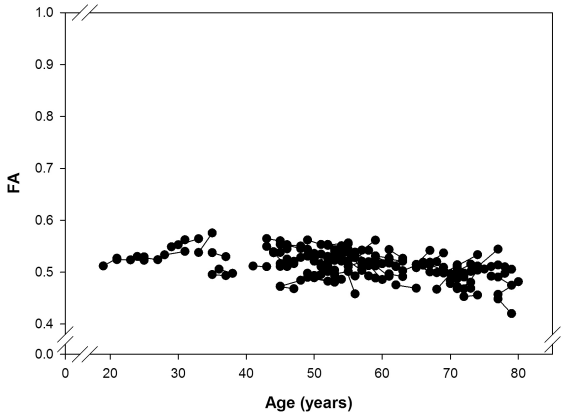
FA Change in Forceps Major with Age



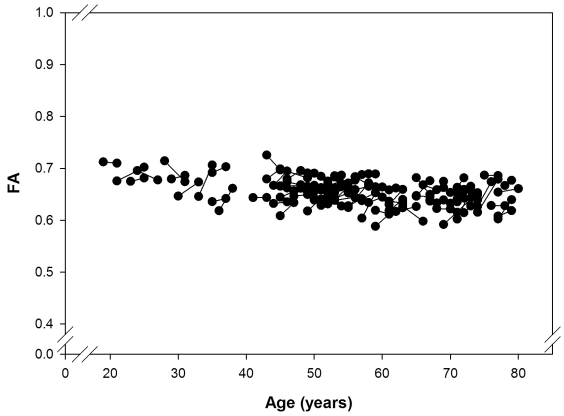
FA Change in Inferior Frontal-Occipital Fasciculus with Age



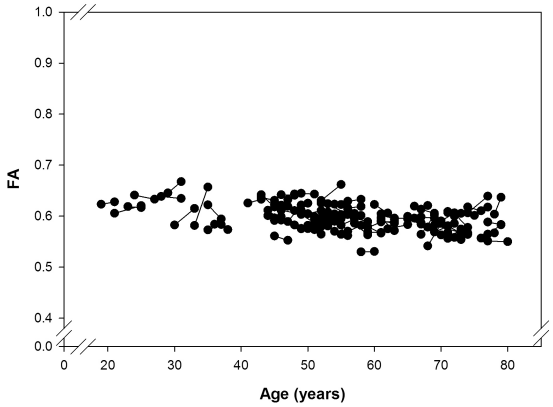
FA Change in Inferior Longitudinal Fasciculus with Age

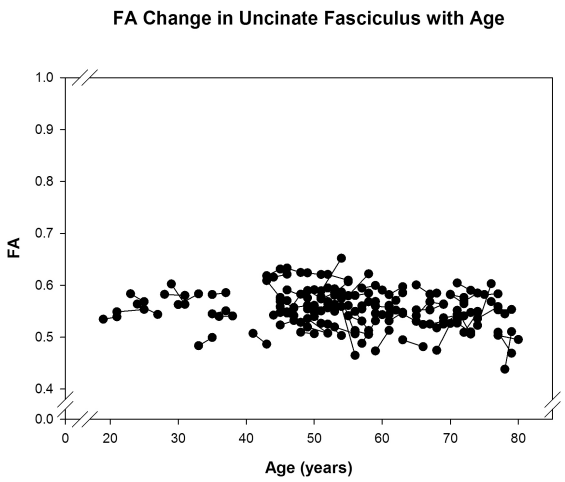


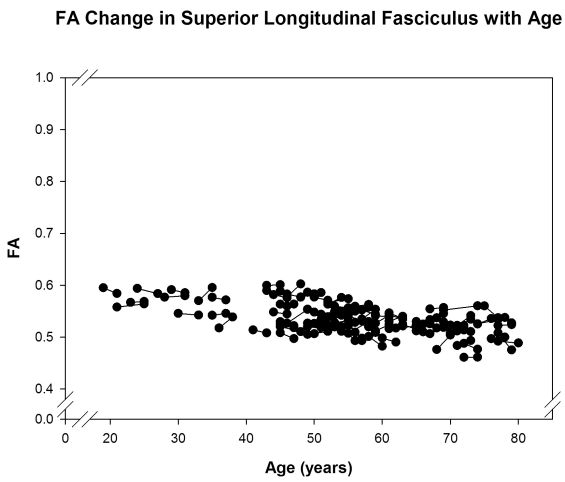
FA Change in Posterior Limb of Internal Capsule vs. Age



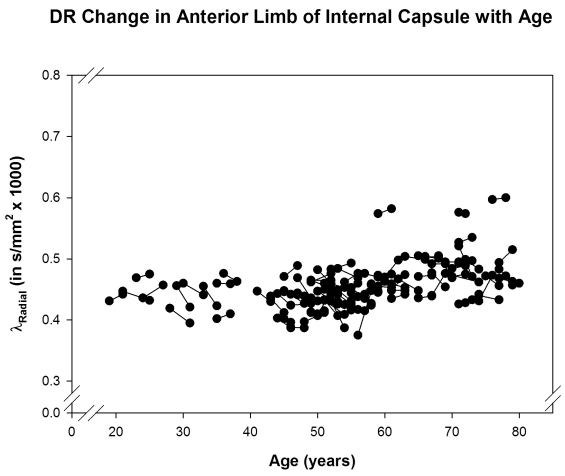
FA Change in Retrolenticular Limb of Internal Capsule vs. Age

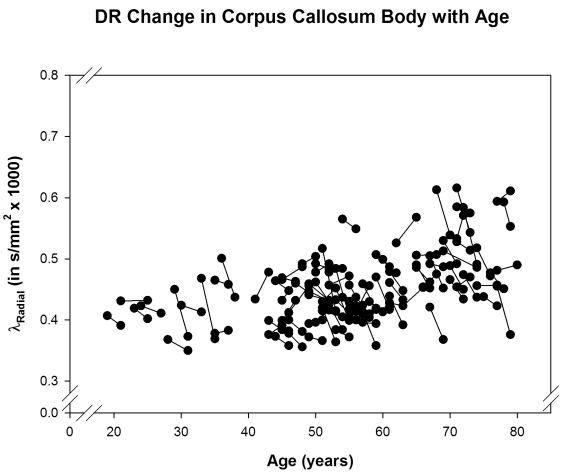


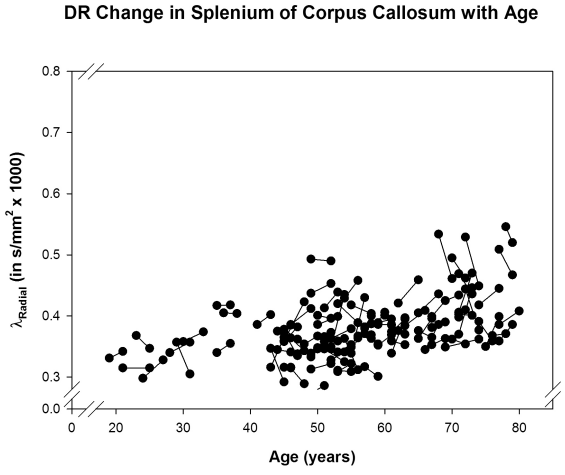


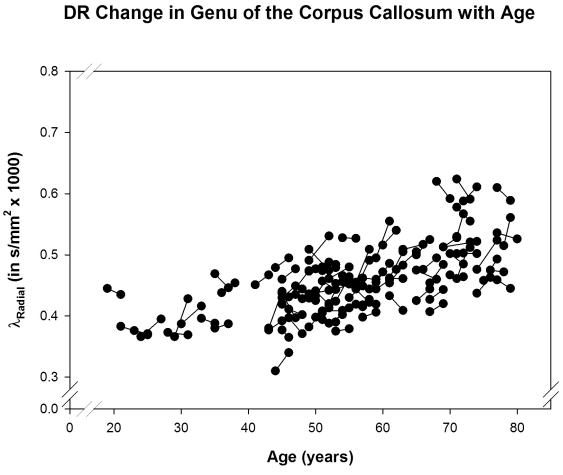


Radial Diffusivity (DR)

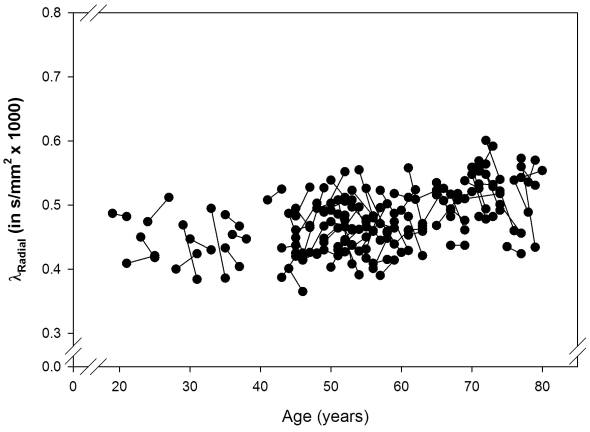




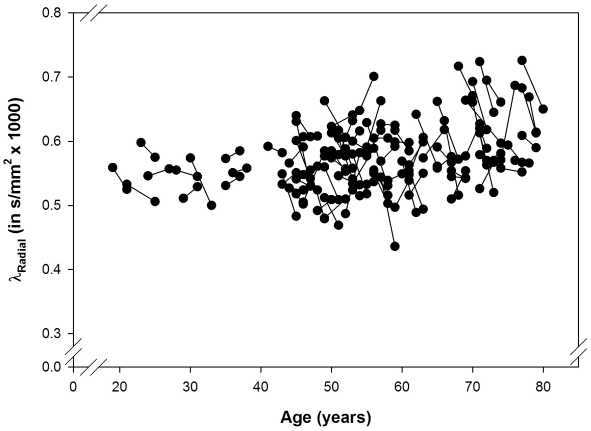




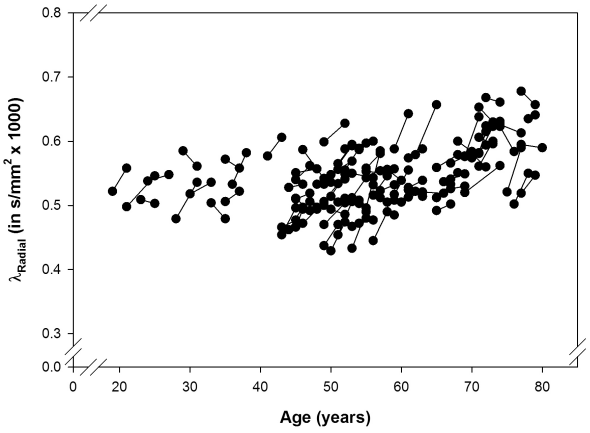
DR Change in Dorsal Cingulum Bundle with Age



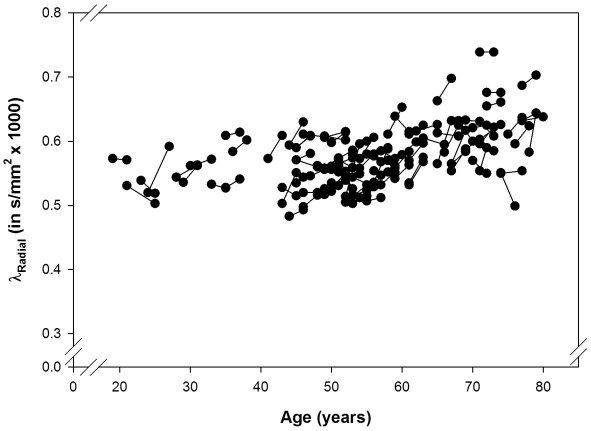
DR Change in Ventral Cingulum Bundle with Age



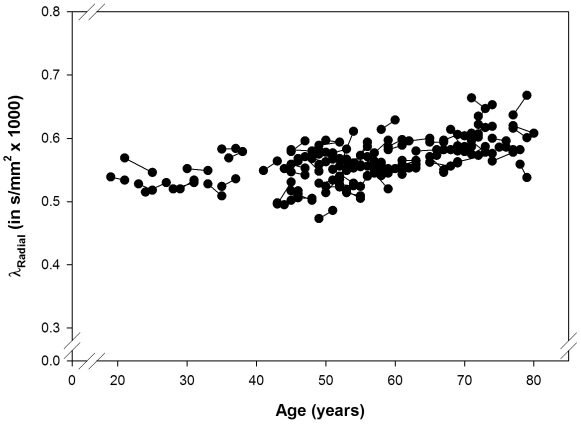
DR Change in Forceps Major with Age



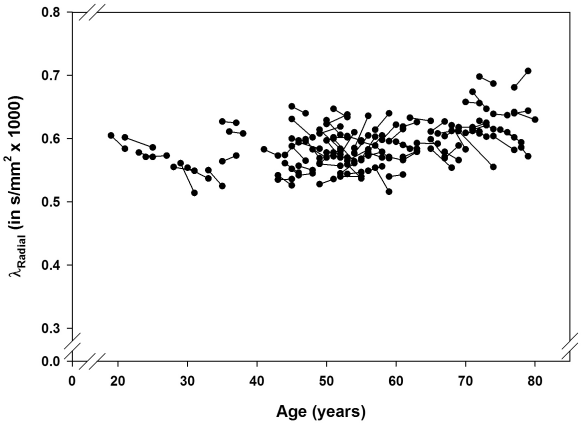
DR Change in Forceps Minor with Age



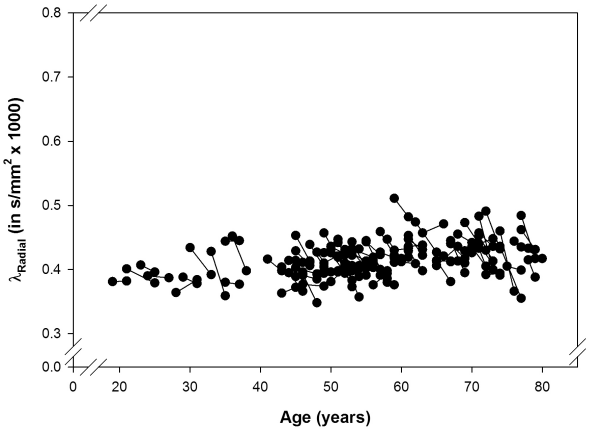
DR Change in Inferior Frontal-Occipital Fasciculus with Age



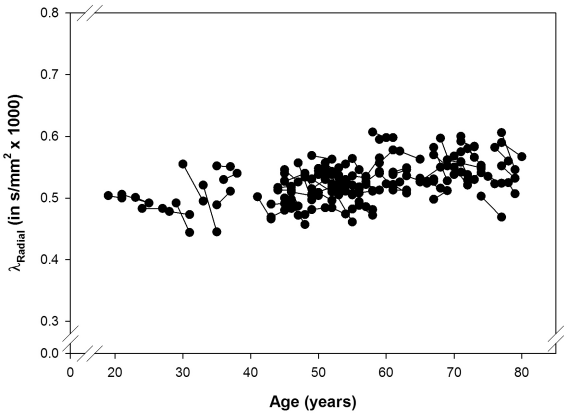
DR Change in Inferior Longitudinal Fasciculus with Age



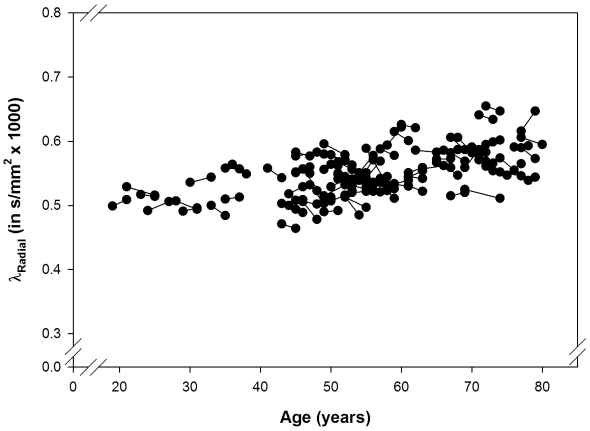
DR Change in Posterior Limb of Internal Capsule vs. Age



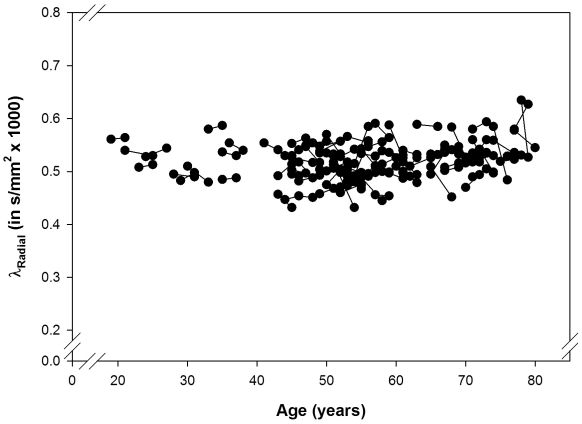
DR Change in Retrolenticular Limb of Internal Capsule vs. Age



DR Change in Superior Longitudinal Fasciculus with Age

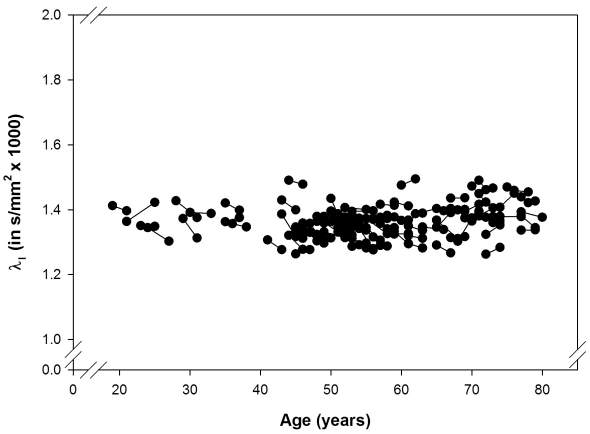


DR Change in Uncinate Fasciculus with Age

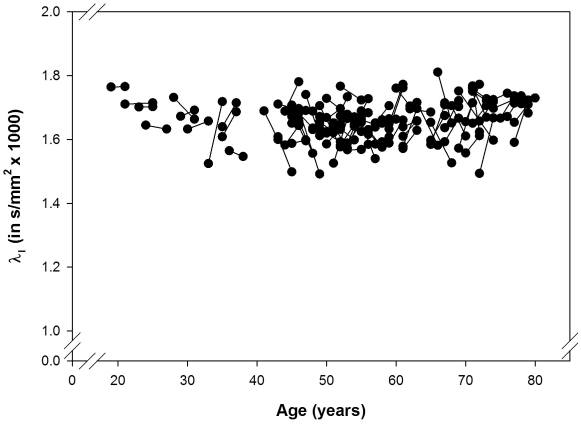


DA

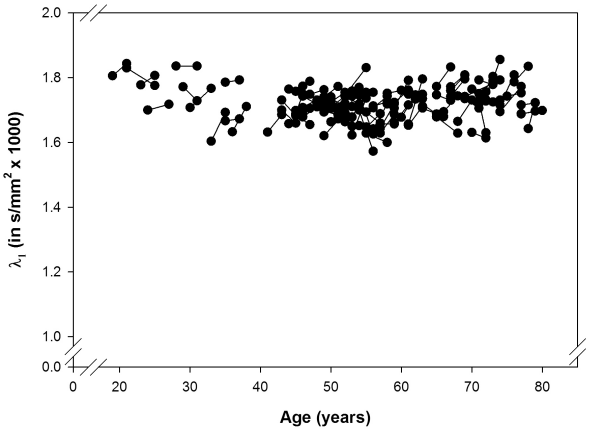
DA Change in Anterior Limb of Internal Capsule with Age



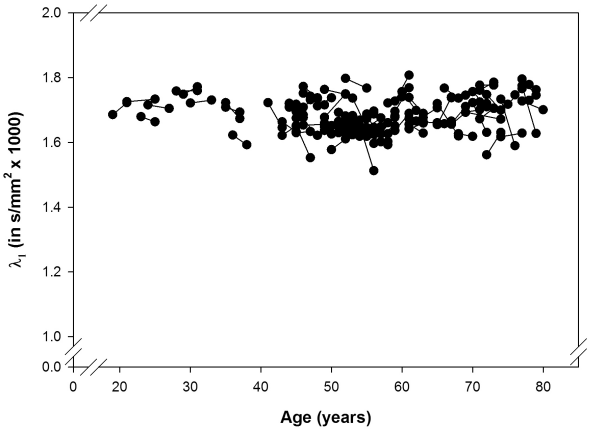
DA Change in Corpus Callosum Body with Age



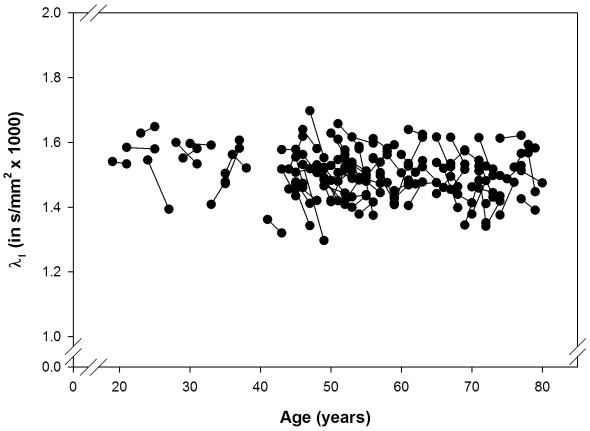
DA Change in Genu of the Corpus Callosum with Age



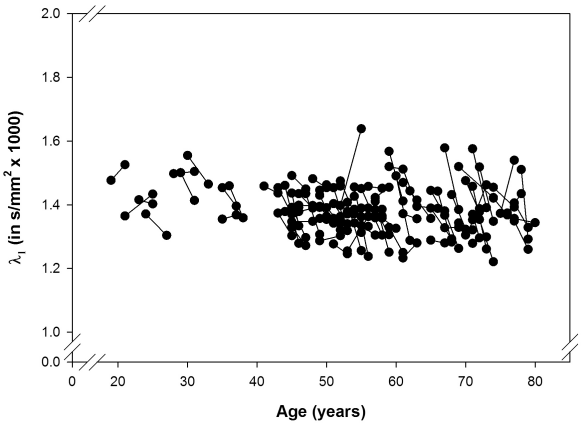
DA Change in Splenium of Corpus Callosum with Age



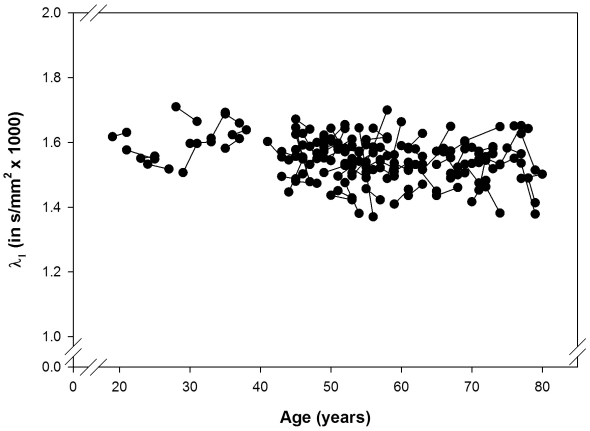
DA Change in Dorsal Cingulum Bundle with Age



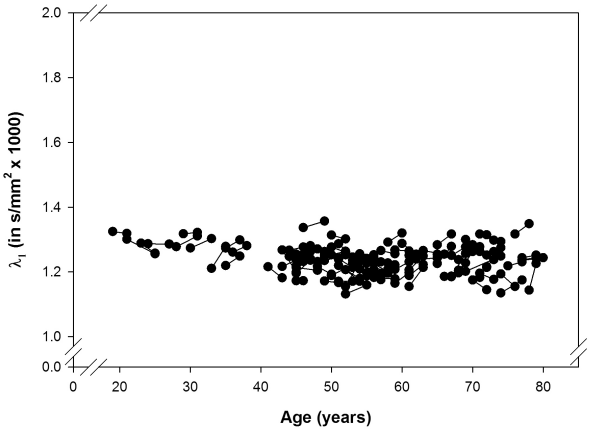
DA Change in Ventral Cingulum Bundle with Age



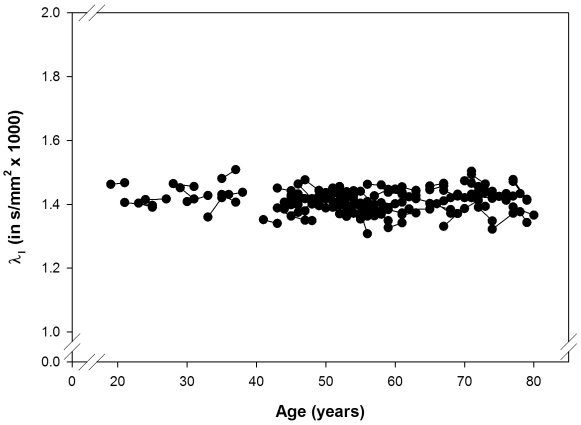
DA Change in Forceps Major with Age



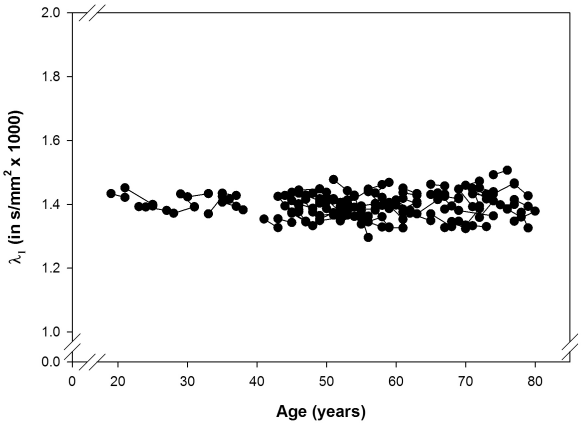
DA Change in Forceps Minor with Age



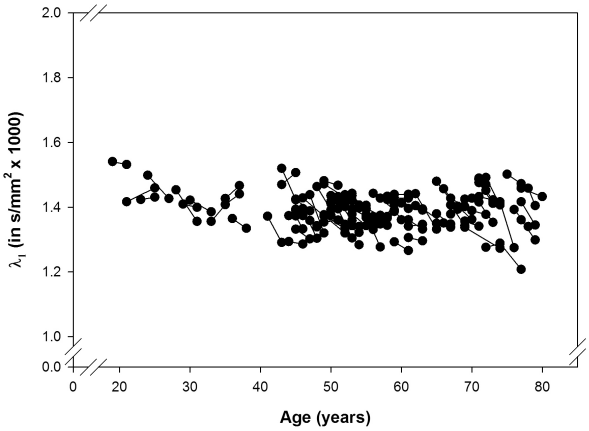
DA Change in Inferior Frontal-Occipital Fasciculus with Age



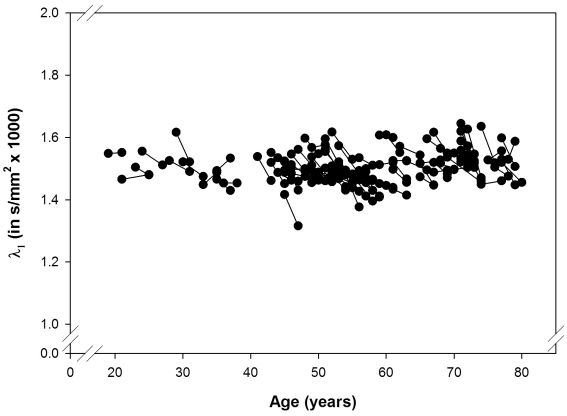
DA Change in Inferior Longitudinal Fasciculus with Age



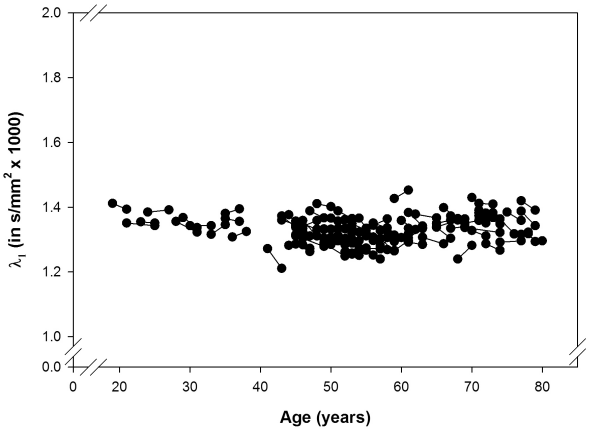
DA Change in Posterior Limb of Internal Capsule vs. Age



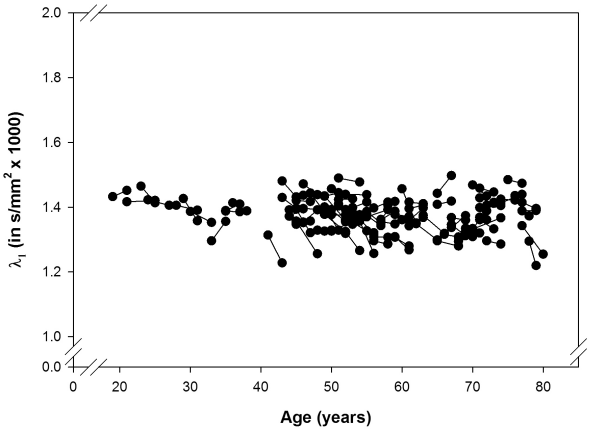
DA Change in Retrolenticular Limb of Internal Capsule vs. Age



DA Change in Superior Longitudinal Fasciculus with Age

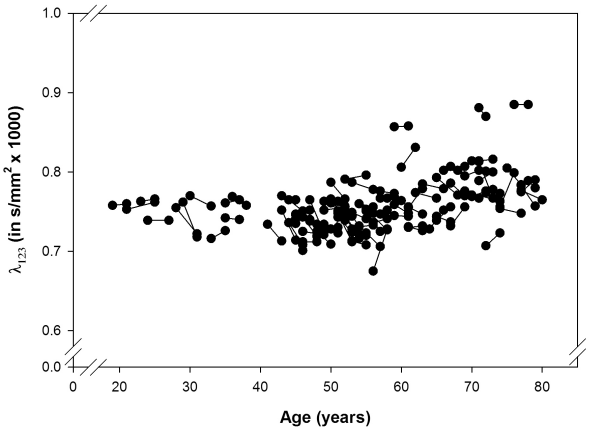


DA Change in Uncinate Fasciculus with Age

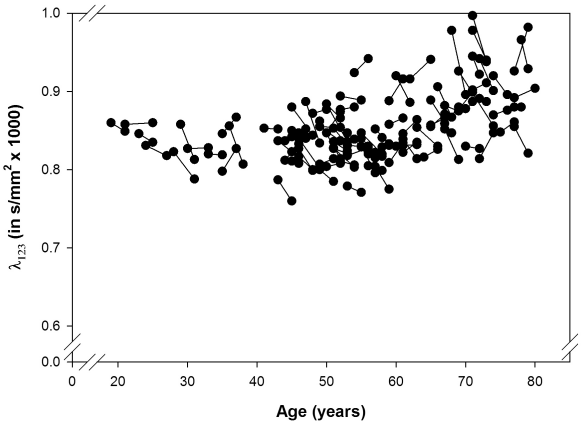


MD

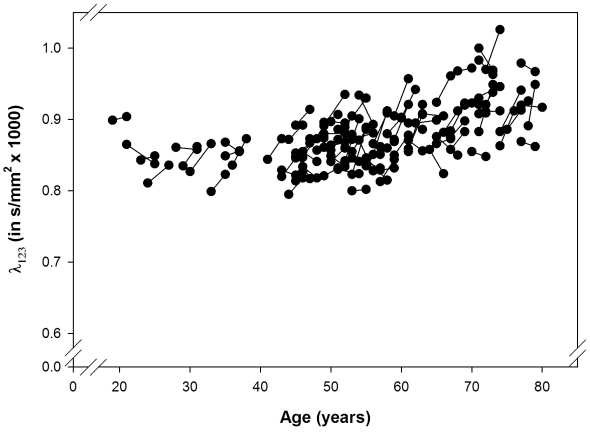
MD Change in Anterior Limb of Internal Capsule with Age



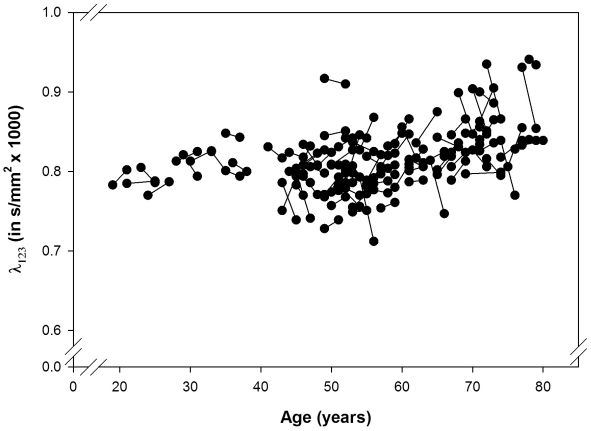
MD Change in Corpus Callosum Body with Age



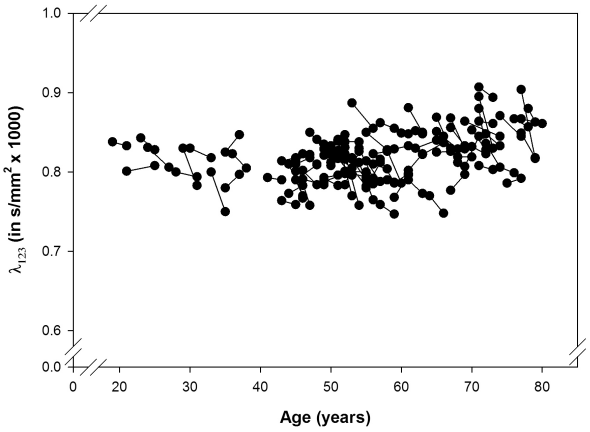
MD Change in Genu of the Corpus Callosum with Age



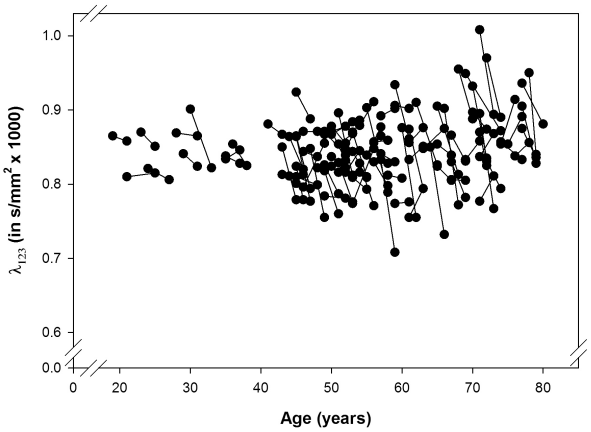
MD Change in Splenium of Corpus Callosum with Age



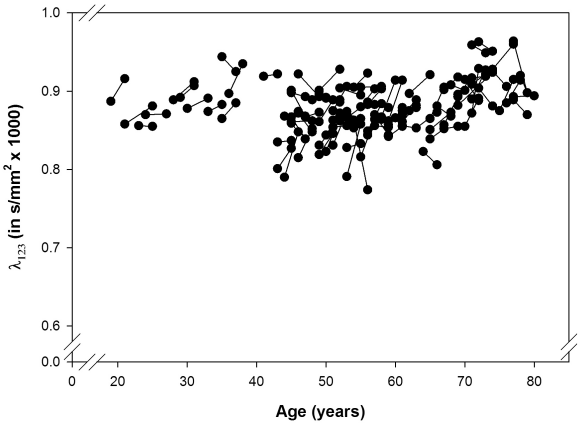
MD Change in Dorsal Cingulum Bundle with Age



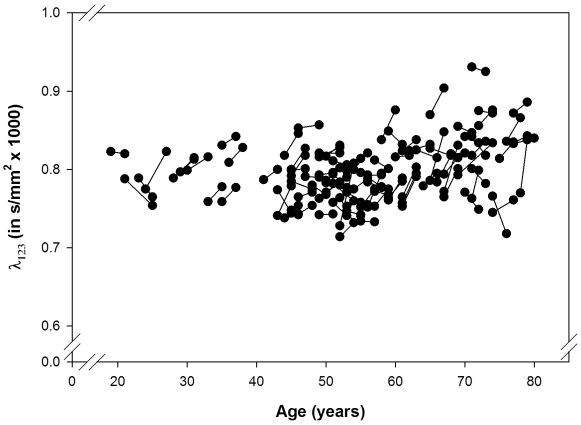
MD Change in Ventral Cingulum Bundle with Age



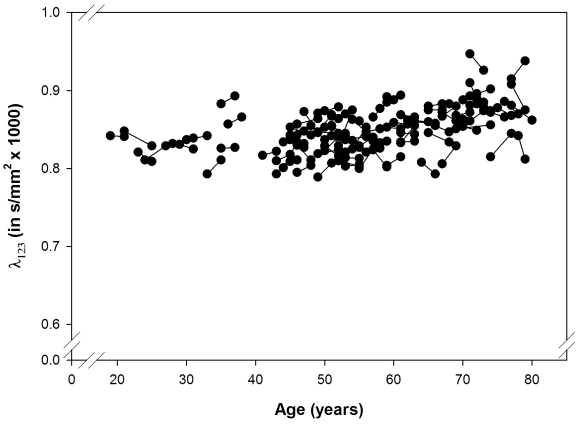
MD Change in Forceps Major with Age



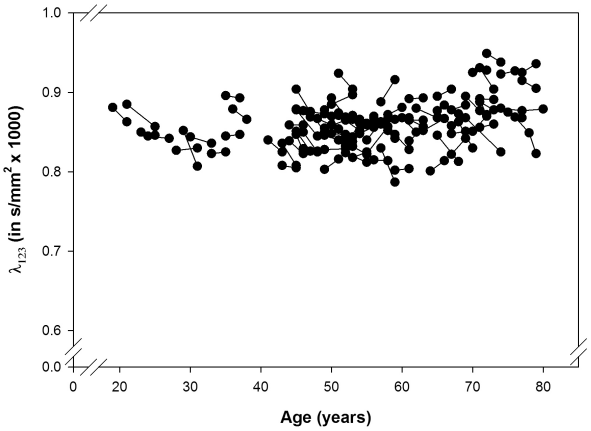
MD Change in Forceps Minor with Age



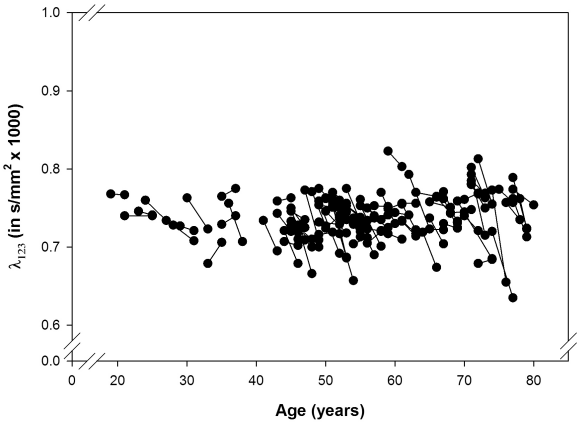
MD Change in Inferior Frontal-Occipital Fasciculus with Age



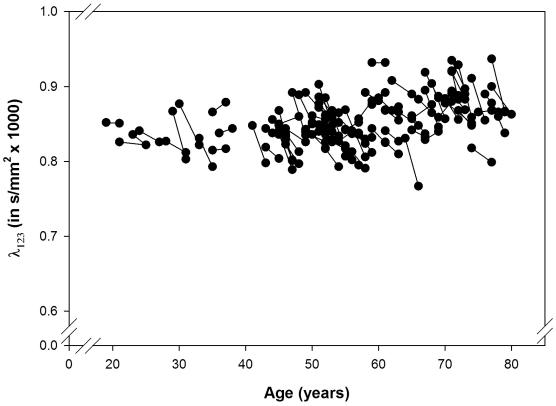
MD Change in Inferior Longitudinal Fasciculus with Age

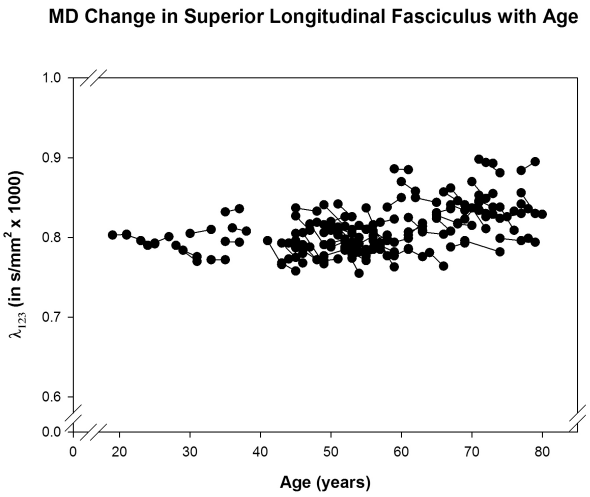


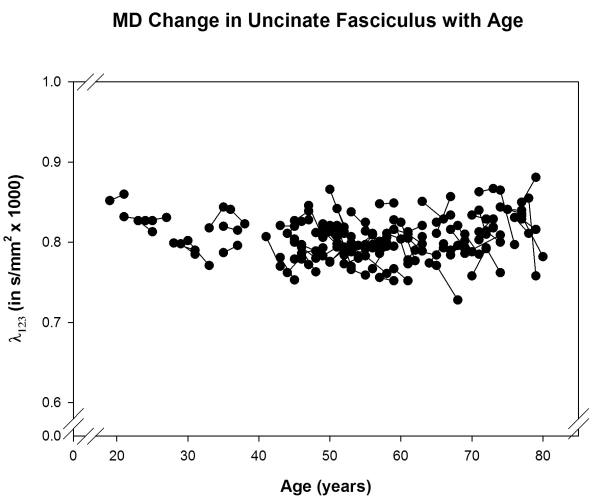
MD Change in Posterior Limb of Internal Capsule vs. Age



MD Change in Retrolenticular Limb of Internal Capsule vs. Age







REFERENCES

Amlien, I. K., Fjell, A. M., Walhovd, K. B., Selnes, P., Stenset, V., Grambaite, R., . . . Fladby, T. (2013). Mild cognitive impairment: cerebrospinal fluid tau biomarker pathologic levels and longitudinal changes in white matter integrity. *Radiology, 266*(1), 295-303.

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., & Buckner, R. L. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron, 56*(5), 924-935.

Barrick, T. R., Charlton, R. A., Clark, C. A., & Markus, H. S. (2010). White matter structural decline in normal ageing: A prospective longitudinal study using tract-based spatial statistics. *NeuroImage, 51*(2), 565-577.

Basser, P. J., & Jones, D. K. (2002). Diffusion-tensor MRI: theory, experimental design and data analysis–a technical review. *NMR in Biomedicine, 15*(7-8), 456-467.

Basser, P. J., & Pierpaoli, C. (1998). A simplified method to measure the diffusion tensor from seven MR images. *Magnetic Resonance in Medicine, 39*(6), 928-934.

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance B, 103*(3), 247-254.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR Biomed, 15*(7-8), 435-455.

Beaulieu, C., & Allen, P. S. (1994). Determinants of anisotropic water diffusion in nerves. *Magnetic Resonance in Medicine, 31*(4), 394-400.

Beaulieu, C., Does, M. D., Snyder, R. E., & Allen, P. S. (1996). Changes in water diffusion due to Wallerian degeneration in peripheral nerve. *Magnetic Resonance in Medicine,*

36(4), 627-631.

Behrens, T., Berg, H. J., Jbabdi, S., Rushworth, M., & Woolrich, M. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage, 34*(1), 144-155.

Bender, A. R., Daugherty, A. M., & Raz, N. (2012). Age and treated hypertension modify change in diffusion properties of white matter in healthy adults. Program No. 826.12. *2012 Neuroscience Meeting Planner*. New Orleans, LA: Society for Neuroscience, 2012. Online.

Bender, A. R., Naveh-Benjamin, M., & Raz, N. (2010). Associative deficit in recognition memory in a lifespan sample of healthy adults. *Psychology and Aging, 25*(4), 940-948.

Bennett, I. J., Madden, D. J., Vaidya, C. J., Howard, J. H., Jr., & Howard, D. V. (2011). White matter integrity correlates of implicit sequence learning in healthy aging. *Neurobiology of Aging, 3*2(12), 2311-2312.

Bereiter C. Some persisting dilemmas in the measurement of change. In: Harris C, editor. Problems in the measurement of change. Madison, WI: University of Wisconsin Press; 1963. pp. 3–20.

Bhagat, Y. A., & Beaulieu, C. (2004). Diffusion anisotropy in subcortical white matter and cortical gray matter: Changes with aging and the role of CSF-suppression. *Journal of Magnetic Resonance Imaging, 20*(2), 216-227.

Brown, W. R., & Thore, C. R. (2011). Review: cerebral microvascular pathology in ageing and neurodegeneration. *Neuropathology and Applied Neurobiology, 37*(1), 56-74.

Buckner, R. L., Kelley, W. M., & Petersen, S. E. (1999). Frontal cortex contributes to human memory formation. *Nature Neuroscience, 2*(4), 311-314.

Bucur, B., Madden, D. J., Spaniol, J., Provenzale, J. M., Cabeza, R., White, L. E., & Huettel, S. A. (2008). Age-related slowing of memory retrieval: contributions of perceptual speed and cerebral white matter integrity. *Neurobiology of Aging*, 29(7), 1070-1079.

Burgmans, S., Gronenschild, E. H., Fandakova, Y., Shing, Y. L., van Boxtel, M. P., Vuurman, E. F., . . . Raz, N. (2011). Age differences in speed of processing are partially mediated by differences in axonal integrity. *NeuroImage, 55*(3), 1287-1297.

Burgmans, S., van Boxtel, M. P., Gronenschild, E. H., Vuurman, E. F., Hofman, P., Uylings, H. B., . . . Raz, N. (2010). Multiple indicators of age-related differences in cerebral white matter and the modifying effects of hypertension. *NeuroImage, 49*(3), 2083-2093.

Burzynska, A. Z., Preuschhof, C., Backman, L., Nyberg, L., Li, S. C., Lindenberger, U., & Heekeren, H. R. (2010). Age-related differences in white matter microstructure: region-specific patterns of diffusivity. *NeuroImage, 49*(3), 2104-2112.

Cattell, R. B., & Cattell, A. (1973). Measuring intelligence with the culture fair tests: Institute for Personality and Ability Testing.

Chao, T.-C., Chou, M.-C., Yang, P., Chung, H.-W., & Wu, M.-T. (2009). Effects of interpolation methods in spatial normalization of diffusion tensor imaging data on group comparison of fractional anisotropy. *Magnetic resonance imaging, 27*(5), 681-690.

Charlton, R. A., Barrick, T. R., McIntyre, D. J., Shen, Y., O'Sullivan, M., Howe, F. A., . . .

Markus, H. S. (2006). White matter damage on diffusion tensor imaging correlates with age-related cognitive decline. *Neurology, 66*(2), 217-222.

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.

Chen, Z. G., T.-Q, L., & Hindmarsh, T. (2001). Diffusion tensor trace mapping in normal adult brain using single-shot EPI technique: A methodological study of the aging brain. *Acta Radiologica, 42*(5), 447-45.

Cherry, K. E., & Park, D. C. (1993). Individual difference and contextual variables influence spatial memory in younger and older adults. *Psychology and Aging, 8*(4), 517-526.

Chiang, M. C., Barysheva, M., Toga, A. W., Medland, S. E., Hansell, N. K., James, M. R., . . . Thompson, P. M. (2011). BDNF gene effects on brain circuitry replicated in 455 twins. *NeuroImage, 55*(2), 448-454.

Clapp, W. C., Rubens, M. T., Sabharwal, J., & Gazzaley, A. (2011). Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proceedings of the National Academy of Sciences, 108*(17), 7212-7217.

Concha, L., Livy, D. J., Beaulieu, C., Wheatley, B. M., & Gross, D. W. (2010). In vivo diffusion tensor imaging and histopathology of the fimbria-fornix in temporal lobe epilepsy. Journal of Neuroscience, 30(3), 996-1002.

Correia, S., Lee, S., Voorn, T., Tate, D., Paul, R., Zhang, S., . . . Laidlaw, D. (2008).

Quantitative tractography metrics of white matter integrity in diffusion-tensor MRI. *NeuroImage, 42*(2), 568-581.

Danielian, L. E., Iwata, N. K., Thomasson, D. M., & Floeter, M. K. (2010). Reliability of fiber tracking measurements in diffusion tensor imaging for longitudinal study. *NeuroImage, 49*(2), 1572-1580.

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Que PASA? The posterior-anterior shift in aging. *Cerebral Cortex, 18*(5), 1201-1209.

de Groot, J. C., Oudkerk, M., Gijn, J., Hofman, A., Jolles, J., & Breteler, M. (2000). Cerebral white matter lesions and cognitive function: the Rotterdam Scan Study. *Annals of Neurology, 47*(2), 145-151.

Della-Maggiore, V., Sekuler, A. B., Grady, C. L., Bennett, P. J., Sekuler, R., & McIntosh, A. R. (2000). Corticolimbic interactions associated with performance on a short-term memory task are modified by age. *Journal of Neuroscience, 2*0(22), 8410-8416.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Hyman, B. T. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage, 31*(3), 968-980.

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory. *Psychology and Aging, 4*(4), 500-503.

Douaud, G., Jbabdi, S., Behrens, T. E., Menke, R. A., Gass, A., Monsch, A. U., . . . Smith, S. (2011). DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *NeuroImage, 55*(3),

880-890.

Dufouil, C., de Kersaint–Gilly, A., Besancon, V., Levy, C., Auffray, E., Brunnereau, L., . . . Tzourio, C. (2001). Longitudinal study of blood pressure and white matter hyperintensities. *Neurology, 56*(7), 921-926.

Eckert, M. A. (2011). Slowing down: age-related neurobiological predictors of processing speed. *Frontiers in Neuroscience*, 5, 25.

Ekstrom, R. B., French, J. W., Harmon, H. H., & Derman, D. (1976). Manual for kit of factor-referenced cognitive tests.

Elwood, R. W. (1991). The Wechsler Memory Scale—Revised: Psychometric characteristics and clinical application. *Neuropsychology Review, 2*(2), 179-201.

Engvig, A., Fjell, A. M., Westlye, L. T., Moberget, T., Sundseth, O., Larsen, V. A., & Walhovd, K. B. (2011). Memory training impacts short-term changes in aging white matter: A longitudinal diffusion tensor imaging study. *Human Brain Mapping, 33*(10), 2390-2406.

Ervin, R. B. (2009). Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States. *National Health Statistics Reports, 13*, 1-7.

Fjell, A. M., Westlye, L. T., Greve, D. N., Fischl, B., Benner, T., van der Kouwe, A. J., . . . Walhovd, K. B. (2008). The relationship between diffusion tensor imaging and volumetry as measures of white matter properties. *NeuroImage, 42*(4), 1654-1668.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of*

Psychiatric Research, 12(3), 189-198.

Ford, J. C., & Hackney, D. B. (1997). Numerical model for calculation of apparent diffusion coefficients (ADC) in permeable cylinders—comparison with measured ADC in spinal cord white matter. *Magnetic Resonance in Medicine, 37*(3), 387-394.

Ford, J. C., Hackney, D. B., Lavi, E., Phillips, M., & Patel, U. (1998). Dependence of apparent diffusion coefficients on axonal spacing, membrane permeability, and diffusion time in spinal cord white matter. *Journal of Magnetic Resonance Imaging, 8*(4), 775-782.

Franceschi, C., Capri, M., Monti, D., Giunta, S., Olivieri, F., Sevini, F., . . . Scurti, M. (2007). Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. *Mechanisms of Ageing and Development, 128*(1), 92-105.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Davidson, K. C., & Thompson, N. M. (1991). Analysis of change: Modeling individual growth. *Journal of Consulting and Clinical Psychology, 59*(1), 27.

Fung, A., Vizcaychipi, M., Lloyd, D., Wan, Y., & Ma, D. (2012). Central nervous system inflammation in disease related conditions: Mechanistic prospects. *Brain Research, 1446*, 144-155.

Geschwind, N. (1965). Disconnexion Syndromes in Animals and Man: I. *Brain, 88*, 237-294.

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage, 51*(3), 943-951.

Gold, B. T., Powell, D. K., Xuan, L., Jiang, Y., & Hardy, P. A. (2007). Speed of lexical decision correlates with diffusion anisotropy in left parietal and frontal white matter: evidence from diffusion tensor imaging. *Neuropsychologia, 45*(11), 2439-2446.

Gons, R. A. R., de Laat, K. F., van Norden, A. G. W., van Oudheusden, L. J. B., van Uden, I. W. M., Norris, D. G., . . . de Leeuw, F. E. (2010). Hypertension and cerebral diffusion tensor imaging in small vessel disease. *Stroke, 41*(12), 2801-2806.

Grady, C. L., McIntosh, A. R., & Craik, F. I. M. (2003). Age-related differences in the functional connectivity of the hippocampus during memory encoding. *Hippocampus, 13*(5), 572-586.

Gunning-Dixon, F. M., Brickman, A. M., Cheng, J. C., & Alexopoulos, G. S. (2009). Aging of cerebral white matter: a review of MRI findings. *International Journal of Geriatric Psychiatry, 24*(2), 109-117.

Gunning-Dixon, F., & Raz, N. (2000). The cognitive correlates of white matter abnormalities in normal aging: A quantitative review. *Neuropsychology, 14*(2), 224-232.

Hannesdottir, K., Nitkunan, A., Charlton, R. A., Barrick, T. R., MacGregor, G. A., & Markus, H. S. (2009). Cognitive impairment and white matter damage in hypertension: A pilot study. *Acta Neurologica Scandinavica, 119*(4), 261-268.

Hasan, K. M., Iftikhar, A., Kamali, A., Kramer, L. A., Ashtari, M., Cirino, P. T., . . . Ewing-Cobbs, L. (2009). Development and aging of the healthy human brain uncinate fasciculus across the lifespan using diffusion tensor tractography. *Brain Research, 1276*, 67-76.

Hasan, K. M., Kamali, A., Iftikhar, A., Kramer, L. A., Papanicolaou, A. C., Fletcher, J. M., &

Ewing-Cobbs, L. (2009). Diffusion tensor tractography quantification of the human corpus callosum fiber pathways across the lifespan. *Brain Research, 1249*, 91-100.

Huang, H., Fan, X., Weiner, M., Martin-Cook, K., Xiao, G., Davis, J., . . . Diaz-Arrastia, R. (2012). Distinctive disruption patterns of white matter tracts in Alzheimer's disease with full diffusion tensor characterization. *Neurobiology of Aging, 33*, 2029-2045.

Hugenschmidt, C. E., Peiffer, A. M., Kraft, R. A., Casanova, R., Deibler, A. R., Burdette, J. H., . . . Laurienti, P. J. (2008). Relating Imaging Indices of White Matter Integrity and Volume in Healthy Older Adults. *Cerebral Cortex, 18*(2), 433-442.

Hultsch, D. F., Hertzog, C., & Dixon, R. A. (1990). Ability correlates of memory performance in adulthood and aging. *Psychology and Aging, 5*(3), 356-368.

Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., & Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebral Cortex, 16*(9), 1349-1360.

Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K., & Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping, 34*(11), 2747-2766.

Jones, D. K., & Cercignani, M. (2010). Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine, 23*(7), 803-820.

Jones, D. K., Knosche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: the do's and don'ts of diffusion MRI. *NeuroImage, 73*, 239-254.

Jöreskog, K. G., & Sörbom, D. (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Lincolnwood, IL: Scientific Software.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9*(4), 637-671.

Kennedy, K. M., & Raz, N. (2009a). Aging white matter and cognition: Differential effects of regional variations in diffusion properties on memory, executive functions, and speed. *Neuropsychologia, 47*(3), 916-927.

Kennedy, K. M., & Raz, N. (2009b). Pattern of normal age-related regional differences in white matter microstructure is modified by vascular risk. *Brain Research, 1297*, 41-56.

Klawiter, E. C., Schmidt, R. E., Trinkaus, K., Liang, H. F., Budde, M. D., Naismith, R. T., . . . Benzinger, T. L. (2011). Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. *NeuroImage, 55*(4), 1454-1460.

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.

Kramer, J. H., Rosen, H. J., Du, A. T., Schuff, N., Hollnagel, C., Weiner, M. W., . . . Delis, D. C. (2005). Dissociations in hippocampal and frontal contributions to episodic memory performance. *Neuropsychology, 19*(6), 799-805

Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *Journal of Neuroscience, 31*(30), 10937-10947.

Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., & Beaulieu, C. (2012). Diffusion tensor imaging of white matter tract evolution over the lifespan. *NeuroImage, 60*(1),

340-352.

Leritz, E. C., Salat, D. H., Milberg, W. P., Williams, V. J., Chapman, C. E., Grande, L. J., . . . McGlinchey, R. E. (2010). Variation in blood pressure is associated with white matter microstructure but not cognition in African Americans. *Neuropsychology, 24*(2), 199-208.

Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: what's change got to do with it? *Psychology and Aging, 2*6(1), 34-47.

Liston, C., Watts, R., Tottenham, N., Davidson, M. C., Niogi, S., Ulug, A. M., & Casey, B. J. (2006). Frontostriatal microstructure modulates efficient recruitment of cognitive control. *Cerebral Cortex, 16*(4), 553-560.

Liu, Y., Spulber, G., Lehtimaki, K. K., Kononen, M., Hallikainen, I., Grohn, H., . . . Soininen, H. (2011). Diffusion tensor imaging and tract-based spatial statistics in Alzheimer's disease and mild cognitive impairment. *Neurobiology of Aging, 32*(9), 1558-1571.

Madden, D. J., Bennett, I. J., Burzynska, A., Potter, G. G., Chen, N. K., & Song, A. W. (2012). Diffusion tensor imaging of cerebral white matter integrity in cognitive aging. *Biochimica et Biophysica Acta, 1822*(3), 386-400.

Madden, D. J., Bennett, I. J., & Song, A. W. (2009). Cerebral White Matter Integrity and Cognitive Aging: Contributions from Diffusion Tensor Imaging. *Neuropsychology Review, 19*(4), 415-435.

Madden, D. J., Spaniol, J., Whiting, W. L., Bucur, B., Provenzale, J. M., Cabeza, R., . . . Huettel, S. A. (2007). Adult age differences in the functional neuroanatomy of visual

attention: a combined fMRI and DTI study. *Neurobiology of Aging, 28*(3), 459-476.

Madden, D. J., Whiting, W. L., Huettel, S. A., White, L. E., MacFall, J. R., & Provenzale, J. M. (2004). Diffusion tensor imaging of adult age differences in cerebral white matter: relation to response time. *NeuroImage, 2*1(3), 1174-1181.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23-44.

McArdle, J. J. (2009). Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology, 60*, 577-605.

McArdle, J. J., & Nesselroade, J. R. (1994). Structuring data to study development and change. In S. H. Cohen & H. W. Reese (Eds.), Life-Span Developmental Psychology: Methodological Innovations (pp. 223-267). Hillsdale, NJ: Erlbaum.

Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29*(2), 177-185.

Metzler-Baddeley, C., Jones, D. K., Belaroussi, B., Aggleton, J. P., & O'Sullivan, M. J. (2011). Frontotemporal connections in episodic memory and aging: a diffusion MRI tractography study. *Journal of Neuroscience, 31*(37), 13236-13245.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Mori, S. (2005). MRI Atlas of Human White Matter. Amsterdam, Netherlands: Elsevier.

Moseley, M., Bammer, R., & Illes, J. (2002). Diffusion-tensor imaging of cognitive performance. *Brain and Cognition, 50*(3), 396-413.

Mueller, R. O. (1996). Basic principles of structural equation modeling: An introduction to LISREL and EQS. New York: Springer Verlag.

Mukherjee, P., Berman, J. I., Chung, S. W., Hess, C. P., & Henry, R. G. (2008a). Diffusion tensor MR imaging and fiber tractography: theoretic underpinnings. *American Journal of Neuroradiology, 29*(4), 632-641.

Mukherjee, P., Chung, S. W., Berman, J. I., Hess, C. P., & Henry, R. G. (2008b). Diffusion tensor MR imaging and fiber tractography: technical considerations. *American Journal of Neuroradiology, 29*(5), 843-852.

Muthén, L.K. and Muthén, B.O. (1998-2012). Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén

Naismith, R. T., Xu, J., Tutlam, N. T., Snyder, A., Benzinger, T., Shimony, J., . . . Song, S. (2009). Disability in optic neuritis correlates with diffusion tensor-derived directional diffusivities. *Neurology, 72*(7), 589.

Naveh-Benjamin, M. (2000). Adult age differences in memory performance: Tests of an associative deficit hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26(5)*, 1170-1187.

Overall, J. E., & Woodward, J. A. (1975). Unreliability of difference scores: A paradox for measurement of change. *Psychological Bulletin, 82*(1), 85-86.

Pantoni, L., & Garcia, J. H. (1997). Pathogenesis of leukoaraiosis: A review. *Stroke, 28*(3), 652-659.

Papagno, C., Miracapillo, C., Casarotti, A., Romero Lauro, L. J., Castellano, A., Falini, A., . . . Bello, L. (2010). What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. *Brain, 134*(2), 405-414.

Penke, L., Munoz Maniega, S., Murray, C., Gow, A. J., Hernandez, M. C., Clayden, J. D., . . .

Deary, I. J. (2010). A general factor of brain white matter integrity predicts information processing speed in healthy older people. *Journal of Neuroscience, 30*(22), 7569-7574.

Pfefferbaum, A., Adalsteinsson, E., & Sullivan, E. V. (2005). Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging. *NeuroImage, 26*(3), 891-899.

Piccinin, A. M., Muniz, G., Sparks, C., & Bontempo, D. E. (2011). An evaluation of analytical approaches for understanding change in cognition in the context of aging and health. *The Journals of Gerontology: Series B, Psychological sciences and social sciences, 66* Suppl 1, i36-49.

Pickering, T. G., Hall, J. E., Appel, L. J., Falkner, B. E., Graves, J., Hill, M. N., . . . Roccella, E. J. (2005). Recommendations for blood pressure measurement in humans and experimental animals. *Hypertension, 45*(1), 142-161.

Pierpaoli, C., Barnett, A., Pajevic, S., Chen, R., Penix, L. R., Virta, A., & Basser, P. (2001). Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *NeuroImage, 13*(6), 1174-1185.

Pierpaoli, C., & Basser, P. J. (1996). Toward a quantitative assessment of diffusion anisotropy. *Magnetic Resonance in Medicine, 36*(6), 893-906.

Pike, N. (2011). Using false discovery rates for multiple comparisons in ecology and evolution. *Methods in Ecology and Evolution, 2*(3), 278-282.

Pitkonen, M., Abo-Ramadan, U., Marinkovic, I., Pedrono, E., Hasan, K. M., Strbian, D., . . . Tatlisumak, T. (2012). Long-term evolution of diffusion tensor indices after

temporary experimental ischemic stroke in rats. *Brain Research, 1445*, 103-110.

Pollack, I., & Norman, D. A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science, 1*, 125-126.

Rabbitt, P., & Lowe, C. (2000). Patterns of cognitive ageing. *Psychological Research, 63*(3), 308-316.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.

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.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology, 12*(1), 95-114.

Raz, N., Lindenberger, U., Ghisletta, P., Rodrigue, K. M., Kennedy, K. M., & Acker, J. D. (2008). Neuroanatomical correlates of fluid intelligence in healthy adults and persons with vascular risk factors. *Cerebral Cortex, 1*8(3), 718-726.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*(11), 1676-1689.

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.

Raz, N., Yang, Y., Dahle, C. L., & Land, S. (2012). Volume of white matter hyperintensities in healthy adults: contribution of age, vascular risk factors, and inflammation-related genetic variants. *Biochim Biophys Acta, 1822*(3), 361-369.

Raz, N., Yang, Y. Q., Rodrigue, K. M., Kennedy, K. M., Lindenberger, U., & Ghisletta, P. (2012). White matter deterioration in 15 months: latent growth curve models in healthy adults. *Neurobiology of Aging, 33*(2), 429.e421-429.e425.

Reuter, M., & Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *NeuroImage, 57*(1), 19-21.

Salat, D. H., Tuch, D. S., Greve, D. N., van der Kouwe, A. J. W., Hevelone, N. D., Zaleta, A. K., . . . Dale, A. M. (2005). Age-related alterations in white matter microstructure measured by diffusion tensor imaging. *Neurobiology of Aging, 26*(8), 1215-1227.

Salthouse, T. A. (1974). Using selective interference to investigate spatial memory representations. *Memory & Cognition, 2*(4), 749-757.

Salthouse, T. A. (1975). Simultaneous processing of verbal and spatial information. *Memory & Cognition, 3*(2), 221-225.

Salthouse, T. A. (1996). General and specific speed mediation of adult age differences in memory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 51*(1), P30-42.

Salthouse, T. A., Fristoe, N., McGuthry, K. E., & Hambrick, D. Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging, 13*(3), 445.

Salthouse, T. A., Hancock, H. E., Meinz, E. J., & Hambrick, D. Z. (1996). Interrelations of age, visual acuity, and cognitive functioning. *The Journals of Gerontology: Series B,*

Social and Psychological Sciences, 51(6), 317-330.

Salthouse, T. A., Kausler, D. H., & Saults, J. S. (1988). Investigation of student status, background variables, and feasibility of standard tasks in cognitive aging research. *Psychology and Aging, 3*(1), 29-37.

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *The Journals of Gerontology: Series B, Social and Psychological Sciences, 50B*(6), 297-306.

Salthouse, T. A., Mitchell, D. R. D., Skovronek, & Babcock, R. L. (1989). Effects of adult age and working memory on reasoning and spatial abilities. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15*(3), 507-516.

Sasson, E., Doniger, G. M., Pasternak, O., & Assaf, Y. (2010). Structural correlates of memory performance with diffusion tensor imaging. *NeuroImage, 50*(3), 1231-1242.

Sasson, E., Doniger, G. M., Pasternak, O., Tarrasch, R., & Assaf, Y. (2012). Structural correlates of cognitive domains in normal aging with diffusion tensor imaging. *Brain Structure and Function, 217*(2), 503-515.

Schmierer, K., Wheeler-Kingshott, C. A., Tozer, D. J., Boulby, P. A., Parkes, H. G., Yousry, T. A., . . . Miller, D. H. (2008). Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation. *Magnetic Resonance in Medicine, 59*(2), 268-277.

Schneider, J. A., & Bennett, D. A. (2010). Where vascular meets neurodegenerative disease. *Stroke, 41*(10 Suppl), S144-146.

Schretlen, D., Pearlson, G. D., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A.,

& Barta, P. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society, 6*(1), 52-61.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*(3), 143-155.

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., . . . Behrens, T. E. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage, 31*(4), 1487-1505.

Song, S. K., Sun, S. W., Ramsbottom, M. J., Chang, C., Russell, J., & Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *NeuroImage, 17*(3), 1429-1436.

Song, S. K., Yoshino, J., Le, T. Q., Lin, S. J., Sun, S. W., Cross, A. H., & Armstrong, R. C. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. *NeuroImage, 26*(1), 132-140

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1), 137-149.

Stebbins, G. T., Carillo, M., Medina, D., deToledo-Morrell, L., Klingberg, T., Poldrack, R. A., . . . Bennett, D. A. (2001). Frontal white matter integrity in aging and its role in reasoning performance: A diffusion tensor imaging study. Program No. 456.3. *2001 Neuroscience Meeting Planner*. San Diego, CA: Society for Neuroscience, 2001.

Stebbins, G. T., Poldrack, R. A., Klingberg, T., Carrillo, M. C., Desmond, J. E., Moseley, M. E., . . . De Gabrieli, J. (2001). Aging effects on white matter integrity and processing

speed: A diffusion tensor imaging study. *Neurology, 56*(8), A374-A375.

Stewart, K. J. (2002). Exercise training and the cardiovascular consequences of type 2 diabetes and hypertension: plausible mechanisms for improving cardiovascular health. *JAMA: The Journal of the American Medical Association, 288*(13), 1622-1631.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643.

Sullivan, E. V., Rohlfing, T., & Pfefferbaum, A. (2010). Longitudinal study of callosal microstructure in the normal adult aging brain using quantitative DTI fiber tracking. *Developmental Neuropsychology, 35*(3), 233-256.

Sullivan, E. V., Rohlfing, T., & Pfefferbaum, A. (2010). Quantitative fiber tracking of lateral and interhemispheric white matter systems in normal aging: Relations to timed performance. *Neurobiology of Aging, 31*(3), 464-481.

Sun, S. W., Liang, H. F., Cross, A. H., & Song, S. K. (2008). Evolving Wallerian degeneration after transient retinal ischemia in mice characterized by diffusion tensor imaging. *NeuroImage, 40*(1), 1-10.

Sun, S. W., Liang, H. F., Le, T. Q., Armstrong, R. C., Cross, A. H., & Song, S. K. (2006). Differential sensitivity of in vivo and ex vivo diffusion tensor imaging to evolving optic nerve injury in mice with retinal ischemia. *NeuroImage, 32*(3), 1195-1204.

Sun, S. W., Neil, J. J., & Song, S. K. (2003). Relative indices of water diffusion anisotropy are equivalent in live and formalin-fixed mouse brains. *Magnetic Resonance in Medicine, 50*(4), 743-748.

Swick, D., & Knight, R. T. (1999). Contributions of prefrontal cortex to recognition memory:

electrophysiological and behavioral evidence. *Neuropsychology, 13*(2), 155-170.

Takao, H., Hayashi, N., Kabasawa, H. and Ohtomo, K. (2012), Effect of scanner in longitudinal diffusion tensor imaging studies. *Human Brain Mapping, 33*, 466–477.

Taki, Y., Thyreau, B., Hashizume, H., Sassa, Y., Takeuchi, H., Wu, K., . . . Kawashima, R. (2013). Linear and curvilinear correlations of brain white matter volume, fractional anisotropy, and mean diffusivity with age using voxel-based and region-of-interest analyses in 246 healthy children. *Human Brain Mapping, 34*(8), 1842-1856.

Teipel, S. J., Meindl, T., Wagner, M., Stieltjes, B., Reuter, S., Hauenstein, K. H., . . . Hampel, H. (2010). Longitudinal changes in fiber tract integrity in healthy aging and mild cognitive impairment: a DTI follow-up study. *Journal of Alzheimer's Disease, 22*(2), 507-522.

Tuch, D. S., Salat, D. H., Wisco, J. J., Zaleta, A. K., Hevelone, N. D., & Rosas, H. D. (2005). Choice reaction time performance correlates with diffusion anisotropy in white matter pathways supporting visuospatial attention. *Proceedings of the National Academy of Sciences, 102*(34), 12212.

Turken, A. U., Whitfield-Gabrieli, S., Bammer, R., Baldo, J. V., Dronkers, N. F., & Gabrieli, J. D. E. (2008). Cognitive processing speed and the structure of white matter pathways: convergent evidence from normal variation and lesion studies. *NeuroImage, 42*(2), 1032-1044.

Turner, R., Le Bihan, D., & Chesnicks, A. S. (1991). Echo-planar imaging of diffusion and perfusion. *Magnetic Resonance in Medicine, 19*(2), 247-253.

Ungvari, Z., Kaley, G., de Cabo, R., Sonntag, W. E., & Csiszar, A. (2010). Mechanisms of

vascular aging: new perspectives. *The Journals of Gerontology: Series A, Biological Science and Medical Science, 65*(10), 1028-1041.

Vernooij, M. W., de Groot, M., van der Lugt, A., Ikram, M. A., Krestin, G. P., Hofman, A., . . . Breteler, M. M. (2008). White matter atrophy and lesion formation explain the loss of structural integrity of white matter in aging. *NeuroImage, 43*(3), 470-477.

Vernooij, M. W., Ikram, M. A., Vrooman, H. A., Wielopolski, P. A., Krestin, G. P., Hofman, A., . . . Breteler, M. M. (2009). White matter microstructural integrity and cognitive function in a general elderly population. *Archives of General Psychiatry, 66*(5), 545-553.

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.

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.

Wakana, S., Jiang, H., Nagae-Poetscher, van Zijl, P. C., & Mori, S. (2003). Fiber tract–based atlas of human white matter anatomy. *Radiology, 230*(1), 77-87.

Wechsler, D. (1987). Wechsler Memory Scale - Revised. New York: Psychological Corporation.

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., . . . Fjell, A. M. (2009). Life-span changes of the human brain white matter: diffusion tensor imaging (dti) and volumetry. *Cerebral Cortex, 20*(9), 2055-2068.

Wheeler-Kingshott, C. A., & Cercignani, M. (2009). About "axial" and "radial" diffusivities.

Magnetic Resonance in Medicine, 61(5), 1255-1260.

Woodcock, R., Johnson, M., & Battery-Revised, W. J. P. E. (1989). Tests of cognitive ability. DLM Teaching Resources, Allen, TX.

Wright, C. B., & Sacco, R. L. (2010). Cardiac index as a correlate of brain volume: Separating the wheat of normal aging from the chaff of vascular cognitive disorders. *Circulation, 122*(7), 676-678.

Zhang, J., Jones, M., DeBoy, C. A., Reich, D. S., Farrell, J. A., Hoffman, P. N., . . . Calabresi, P. A. (2009). Diffusion tensor magnetic resonance imaging of Wallerian degeneration in rat spinal cord after dorsal root axotomy. *Journal of Neurosc*ience*, 29*(10), 3160-3171.

Ziegler, D. A., Piguet, O., Salat, D. H., Prince, K., Connally, E., & Corkin, S. (2010). Cognition in healthy aging is related to regional white matter integrity, but not cortical thickness. *Neurobiology of Aging, 31*(11), 1912-1926.

ABSTRACT

CHANGES IN CEREBRAL WHITE MATTER, VASCULAR RISK AND COGNITION ACROSS THE ADULT LIFESPAN

by

ANDREW R. BENDER

May 2014

Advisor: Dr. Naftali Raz

Major: Psychology (Behavioral and Cognitive Neuroscience)

Degree: Doctor of Philosophy

Numerous studies over the past decade have used diffusion tensor imaging (DTI) to examine the associations between age, diffusion and anisotropy measures of cerebral white matter (WM), and cognitive performance. However, few have examined relationships between intra-individual change in DTI measures of WM and cognitive function. It is possible that the extant cross-sectional findings are a poor representation of age-related change in WM and cognition. The present study used latent difference-score modeling (LDM) to assess change over two years in DTI indices fractional anisotropy (FA), radial diffusivity (DR), axial diffusivity (DA) and mean diffusivity (MD). In addition, we examined the effects of WM change on concomitant change in age-sensitive cognitive domains, while controlling for individual differences in duration of hypertension treatment, a common marker of vascular risk. A sample of 96 healthy participants spanning the adult lifespan underwent DTI scanning and cognitive testing at baseline and following a two-year delay. Univariate latent difference score model (LDM) analyses showed cross-sectional associations between DTI measures and age did not accurately describe actual change in WM. Regions including the

genu of the corpus callosum, forceps major, and forceps minor showed reliable reductions in FA and increases in MD. In addition, change differed between DR and DA, suggesting the importance of evaluating these measures separately, rather than relying solely on MD as an index of diffusivity. Comparison of change in DTI measures and memory revealed baseline individual differences in FA, DA, and DR in dorsal and ventral cingulum bundle predicted two-year change in associative memory, while baseline DR and FA in multiple regions predicted change in recognition performance. The present findings demonstrate the necessity of longitudinal evaluation of change in WM and cognition.

AUTOBIOGRAPHICAL STATEMENT

I received my Bachelor of Science degree in Psychology from the University of Arizona in 1998. Between 1999 and 2003, I worked in the private sector as a project manager, programmer, graphic designer, and research analyst. I subsequently returned to academia and earned a Master's of Arts degree in Psychology from San Diego State University in 2006. In 2005 I joined the doctoral program in Psychology at Wayne State University, concentrating in the area of Behavioral and Cognitive Neuroscience. Concurrently, I entered the pre-doctoral training program at the Wayne State University, Institute of Gerontology. While at Wayne State University, I served as Treasurer and President of the Student Neuroscience Society, Vice President and President of the Institute of Gerontology Graduate Student Organization, and Graduate Student Representative to the American Psychological Association's Division on Aging and Adult Development. I am the recipient of numerous awards, including the Thomas C. Rumble Graduate Fellowship, the Duncan McCarthy Award from the Michigan Chapter of the Society for Neuroscience, the Wayne State University Campus Life Leadership Award, the Julie A. Thomas Endowed Scholarship in Psychology, and numerous travel awards from the Institute of Gerontology and the Wayne State University Dean of Students. I am first author of five publications in international journals including Neuropsychologia, Neuropsychology, Psychology and Aging, and the Journal of Cognitive Neuroscience. In addition, I have presented my work at international conferences including the Society for Neuroscience, Gerontological Society of America, Cognitive Neuroscience Society, and Cognitive Aging Conference. This dissertation, "Changes in cerebral white matter, vascular risk and cognition across the adult lifespan," was supervised by Dr. Naftali Raz.