University of California

Los Angeles

Development of human brain connectivity in health and disease

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Biomedical Engineering

by

John Benjamin Colby

2012

Abstract of the Dissertation

Development of human brain connectivity in health and disease

by

John Benjamin Colby

Doctor of Philosophy in Biomedical Engineering

University of California, Los Angeles, 2012

Professor Elizabeth R. Sowell, Co-chair

Professor Mark S. Cohen, Co-chair

White matter development in the human brain undergoes a uniquely extended developmental trajectory, and the maturation of this complex network of connections is broadly relatable to real-world measures of cognitive ability. We begin this dissertation with a review of the literature on structural brain development, in order to provide a useful background layer for our discussion (Chapter 1). Next, we report a case study investigating white matter abnormalities in the context of prenatal methamphetamine exposure, and their cognitive correlates (Chapter 2). Then we describe our successfully-funded NIH grant proposal to 1) develop novel image analysis techniques for investigating white matter connectivity, and 2) apply them to the study of a) typical frontal lobe white matter maturation, b) its relation to executive functioning, and c) how these processes are affected by prenatal alcohol exposure (Chapter 3). The results of this effort are described in Chapter 4, Chapter 5, and Chapter 6. Finally, we supplement the work on our UCLA cohort with exciting results from two large, multi-site, collaborative efforts: 1) ADHD–200, an initiative to employ brain mapping findings in a machine learning environment for the diagnostic classification of individual subjects with attention deficit hyperactivity disorder (ADHD) (Chapter 7), and 2) PING (the Pediatric Imaging, Neurocognition, and Genetics study), where we are leading an effort to provide the most comprehensive mapping to-date of typical white matter development (Chapter 8).

The dissertation of John Benjamin Colby is approved.

Nader Pouratian

Paul M. Thompson

Mark S. Cohen, Committee Co-chair

Elizabeth R. Sowell, Committee Co-chair

University of California, Los Angeles

2012

*While **science** deals with fundamental questions about our world and their answers, **engineering** addresses practical problems and their solutions. Nowhere is the synergism between these approaches clearer than as it relates to their application in **medicine**. By focusing increasing effort at this junction, we can most effectively translate our research advancements into practical clinical progress towards the ultimate goal of alleviating human suffering and disease.*

Table of Contents

List of Figures

List of Tables

Acknowledgments

My training has been supported by the following grants: Medical Scientist Training Program (National Institute of General Medical Sciences (NIGMS), T32 GM008042), Ruth L. Kirschstein National Research Service Award (National Institute on Alcohol Abuse and Alcoholism (NIAAA), F30 AA020431), NeuroImaging Training Program (National Institute on Drug Abuse (NIDA), R90 DA023422).

Chapter 1 is a version of: Colby J. B., O'Hare E. D., and Sowell E. R. Brain development: Birth through adolescence. In: Comprehensive Developmental Neuroscience. Elsevier, In press.

Chapter 2 is a version of: Colby J. B., Smith L., O'Connor M. J., Bookheimer S. Y., Van Horn J. D., and Sowell E. R. White matter microstructural alterations in children with prenatal methamphetamine/polydrug exposure, Submitted.

Chapter 4 is a reprint of: Colby J. B., Van Horn J. D., and Sowell E. R. (2011). Quantitative in vivo evidence for broad regional gradients in the timing of white matter maturation during adolescence. Neuroimage 54(1), 25–31.

Chapter 5 is a reprint of: Colby J. B., Soderberg L., Lebel C., Dinov I. D., Thompson P. M., and Sowell E. R. (2012). Along-tract statistics allow for enhanced tractography analysis. Neuroimage 59(4), 3227–3242.

Chapter 7 is a collaboration with Jesse Brown, Pamela Douglas, Jeffrey Rudie, and Zarrar Shehzad.

Vita

2003–2004	Laboratory Assistant, Neurocrine Biosciences, Inc..
2004	Undergraduate researcher, Distributed Systems Technology Center (Abramson Lab), Monash University, VIC, AU.
2004–2005	Undergraduate researcher, Cardiac Mechanics Research Group (McCulloch Lab), Department of Bioengineering, UC San Diego.
2005	B.S. (Bioengineering: Premedical) and B.S. (Molecular Biology), UC San Diego.
2005–2006	Engineer/Scientist, Science Applications International Corp. (SAIC).
2011	M.S. (Biomedical Engineering), UCLA.

Publications

Colby J. B., Soderberg L., Lebel C., Dinov I. D., Thompson P. M., and Sowell E. R. (2012). Along-tract statistics allow for enhanced tractography analysis. Neuroimage 59(4), 3227–3242.

Colby J. B., Van Horn J. D., and Sowell E. R. (2011). Quantitative in vivo evidence for broad regional gradients in the timing of white matter maturation during adolescence. Neuroimage 54(1), 25–31.

Colby J. B., Smith L., O'Connor M. J., Bookheimer S. Y., Van Horn J. D., and Sowell E. R. The

effects of prenatal methamphetamine/polydrug exposure on white matter microstructure, Submitted.

Lebel C., Warner T., Colby J. B., Soderberg L., Roussotte F., Behnke M., Eyler F. D., and Sowell E. R. Imaging white matter microstructure in children with prenatal cocaine exposure, Submitted.

Colby J. B., O'Hare E. D., and Sowell E. R. Brain development: Birth through adolescence. In: Comprehensive Developmental Neuroscience. Elsevier, In press.

CHAPTER 1

Structural brain development review

1.1 Abstract

This discussion will cover the major features of human structural brain development during the period from birth through adolescence, as viewed with magnetic resonance imaging (MRI), and the ways in which these phenomena relate to concurrent cognitive advancements. A special focus will be put on the prominent sex-specific, regional, and temporal variations that characterize this dynamic process.

1.2 Introduction

Human brain development is a dynamic process that begins *in utero* and continues prominently through childhood, adolescence, and young adulthood. While strongly influenced by genetic factors, the environment also prominently affects brain maturation by acting on the cellular and macroscopic levels. This experiential learning impacts both brain structure and function through forms of neuronal plasticity that continue throughout our lifetimes. However, despite the fact that investigating brain development is undoubtedly one of the keys to appreciating how we emerge as unique human beings and how this process can go awry in disease, our understanding of this important period has historically been hindered by two main factors. First, there has been a lack of reliable postmortem data, as thankfully children are generally healthy during development. Second, technological limitations of past methods like positron emission tomography (PET) and computed tomography (CT) often imposed some modest risk of harm to the subject (*e.g.* ionizing radiation),

which made the study of healthy typically-developing children ethically questionable. The situation changed dramatically with the dissemination of magnetic resonance imaging (MRI) technology during the 1980s, which not only offers higher quality images of the brain parenchyma than ultrasound, X-Ray, CT, or PET, but also does so in a way that is remarkably safe for the subject.

This discussion will begin with a review of the historical postmortem and histological literature, and will then move on to the groundbreaking neuroimaging investigations of the 1990s that first examined brain development with this MRI technology. A collection of more detailed phenomena will then be examined, which have been uncovered with advanced brain mapping techniques, and have come together as a set of classic features that characterize typical brain development. Finally, we will conclude with a discussion of the cutting-edge efforts being made to integrate these diverse observations within a more generalized "multimodal" imaging framework, and to relate them to advancements in cognitive development. A focus on prominent sex-specific, regional, and temporal variations will be continually threaded throughout this discussion.

1.3 Postmortem studies and histology

1.3.1 Comparison to MRI

Although datasets were sparse, postmortem and histological studies were able to provide key insight into normal brain structure and development, as well as pathology, decades before the introduction of neuroimaging methods like PET and MRI. Further, the rich literature that developed from this early effort has provided a strong foundation of data against which newer imaging modalities can be validated. Compared to a modern imaging method like MRI, there are several distinct advantages and disadvantages of these postmortem studies. Not only are datasets relatively small in postmortem samples, as mentioned above, but longitudinal studies — valued for their statistical power to detect changes over time within individuals among the highly variable population — are impossible to conduct. Conversely, because postmortem methods can directly visualize the brain tissue, spatial resolution far exceeds even the best neuroimaging protocol, and there is less

validation needed to ensure that the raw signal being measured faithfully represents the underlying neuronal architecture. Artifacts, though, are an important concern for either method. While postmortem methods may introduce artifacts due to cell death, fixation/staining procedures, and morphological changes due to osmotic pressure and mechanical damage, MRI data suffer artifacts from other sources like magnetic susceptibility effects (signal loss in regions near large caverns of air), local image distortions caused by magnetic field inhomogeneities, and partial volume effects that occur when different structures fall within the same voxel. Many of these issues present less of a problem for the interpretation of larger data sets obtained with MRI, relative to postmortem data, as effects of artifacts generally become small as the number of samples becomes large. However, it should still be remembered that MRI only offers a wide-angle indirect view of tissue, which cannot reach down to the cellular level, and must be observed through the complex lens of magnetic resonance.

1.3.2 Synaptogenesis and pruning

By the time an infant is born, the human brain already contains on the order of 100 billion neurons (Kandel *et al.*, 2000). The period of rapid overall brain growth that began *in utero* continues after birth through the first years of life. Surprisingly, however, postmortem studies during the early part of the 20th century showed that total brain volume and weight actually plateau early and reach approximately 90% of their adult values by age 5 (Dekaban, 1978; Riddle *et al.*, 2010; Vignaud, 1966).

Even during this early period of pronounced overall growth, brain development is characterized as a dynamic process with both progressive and regressive changes that are influenced by complex genetic influences as well as experience-dependent plasticity due to environmental influences. As the infant brain grows in size, it also grows in complexity. Neurons undergo dendritic branching, forming an arbor of neural connections through synaptogenesis, and then ultimately refine this global brain network though the processes of myelination and synaptic pruning. Much of our understanding of the complex balance between synaptogenesis and synaptic pruning has evolved from

the seminal histological work performed by Huttenlocher and colleagues, who mapped synaptic density in different areas of the brain throughout childhood. Overall synaptic density is comparable to the adult level at birth. It then rises even further through the first year of life to its peak at 12–18 months, and then decreases during late childhood and young adulthood towards a stable adult plateau of 1 billion synapses/mm^3 (Huttenlocher, 1979). This has helped to form the theory that the flexible groundwork laid through an initial overabundance of connections gives way to a reduced — but more targeted and efficient — network through experience-dependent synaptic pruning. Interesting regional variations were also observed during these studies, with primary visual and auditory cortex reaching their peak synaptic densities earlier than prefrontal cortex (Huttenlocher, 1979; Huttenlocher and Dabholkar, 1997; Huttenlocher *et al.*, 1982). The extended period of synapse elimination also has regional variations, with pruning ending by age 12 in the auditory cortex but continuing through mid-adolescence in the prefrontal cortex. This temporal pattern parallels concurrent gains in the cognitive domains that are thought to relate to these regions (Luna *et al.*, 2004; Spear, 2000).

1.3.3 Myelination

Myelination of axonal projections by oligodendroglia is also a prominent component of early brain development. This process begins *in utero*, continues rapidly through the first 5 years of life, and remarkably extends — although at a slower rate — through young adulthood. Intracortical histological preparations by Kaes in 1907 were some of the first to demonstrate this prolonged trajectory of myelination, and also its striking regional variability in timing (Kaes, 1907; Kemper, 1994). These slides not only demonstrated earlier trajectories in some areas (posterior temporal, pre-central, and post-central cortex) than others (superior parietal, anterior temporal, anterior frontal cortex), but also showed that regions with a more protracted developmental trajectory have more pronounced changes during older age. This has helped to form the "first-in-last-out" theory of aging (Davis *et al.*, 2009), which suggests that higher-order cognitive manifestations (*e.g.* problem solving and logical reasoning) — some of the last to develop (Luna *et al.*, 2004) — are some of the first to de-

generate in old age. Furthermore, the visible spread of myelin outwards into the cortex results in an apparent cortical thinning, which suggests that normal developmental decreases in cortical thickness (discussed later) may be due, in part, to this progressive increase in myelin, and not simply to regressive changes like synaptic pruning and cell loss.

These initial observations in intracortical tissue were extended to the white matter in pioneering work performed by Yakovlev and Lecours in the 1960s. They demonstrated that white matter myelination begins *in utero* during the second trimester of pregnancy, and continues throughout young adulthood (Yakovlev and Lecours, 1967). Additionally, they extended the earlier observations of regional variations in the timing of myelination, and described a general posterior-to-anterior trend in the timing of white matter myelination during development that has also been replicated in other samples (Kinney 1988). Later independent research targeting the hippocampal formation has also noted striking increases in myelination, with a 95% increase observed in the extent of myelination relative to brain weight during the first two decades of life. Surprisingly, the authors noted that expanding myelination continued even through the fourth to sixth decades of life (Benes *et al.*, 1994). Taken together, these observations suggest that structural white matter development, in the form of advancing myelination, proceeds in tune with overall cognitive development — with areas involved in lower-order sensory and motor function myelinating earlier than areas involved with higher-order executive function. This correlated timing implies there may be some relationship between advancing brain function and increased myelination; however, postmortem studies are limited from investigating this directly.

1.3.4 Sex differences

A pronounced sexual dimorphism in overall brain size emerges during the first 5 years of human brain development, with males having brains that are, on average, approximately 10% larger than females at their adult plateau (Dekaban, 1978). This simple and widely reproducible observation has served as a catalyst for continued interest in the study of sex-specific differences during brain development in order to 1) map other detailed components of brain development that may also

show sex-specific differences, 2) determine if there are any cognitive correlates with these findings (Kimura, 1996), 3) establish what — if not total volume of brain matter — are the driving structural contributors to individual cognitive differences in areas like language skills and overall intelligence, and, perhaps most importantly, 4) better understand and clinically address the range of neuropsychiatric disorders that tend to emerge during adolescence with prominent sex-specific affinities (Marsh *et al.*, 2008). Interestingly, while some of this sex-specific variance in brain size can be attributed to height, which is consistent with broader trends across different mammalian species, there remains a significant sex-specific effect on brain size even when differences in body size are taken into account (Peters *et al.*, 1998). Although the brains of adult males tend to be larger than adult females, this increase is actually smaller than what would be predicted based on differences in adult height alone. Histological findings indicating a 15% higher neuronal density in males than females are consistent with this (Rabinowicz *et al.*, 2002), although conflicting reports from other studies prohibit firm conclusions on this point (Haug, 1987; Pakkenberg and Gundersen, 1997). A consideration of the fact that females actually tend to be taller than males during late childhood, perhaps due to faster pubertal maturation in girls, further weakens the idea of such a simple allometric relationship when age-matched males and females are compared (Giedd *et al.*, 2006). These discrepancies highlight the diversity that exists among the postmortem literature on the topic of sex differences in brain development, which is also likely to be influenced by a variety of confounds (including cohort effects and observational bias) that have made interpretation challenging (Peters *et al.*, 1998). Additionally, these reports are limited to either simple global measures like total brain volume or weight, or very local measures like neuronal density, and generally do not account for regional variations in measures like cortical thickness and folding complexity (Luders *et al.*, 2004; Sowell *et al.*, 2007).

1.3.5 Summary

The central theme that emerges from this early postmortem work is that brain development from birth through adolescence is a uniquely dynamic process, encompassing both progressive and re-

gressive events, with varying magnitudes and timing across different regions of the brain. In particular, the concurrent decrease in synaptic density and increase in white matter myelination is consistent with the principle of selective specialization, which has been postulated to be the driving force behind the creation of cognitive networks and thought to form the foundation for higher cognitive processes (Fuster, 2002; Post and Weiss, 1997; Tsujimoto, 2008). The initial overabundance of neurons and synapses during infancy is thought to provide a flexible substrate through which activity-dependent plasticity can fine-tune neural network activity, via processes like synaptic pruning, which continue robustly through adolescence and in some form throughout life.

1.4 *In vivo* volume analyses

With the development of MRI, not only were clinicians provided with a superior technology for the diagnosis of brain injury and disease (Barkovich, 2006; Panigrahy and Blüml, 2009; Prager and Roychowdhury, 2007), but researchers were also provided with an unparalleled technology for the study of typical brain development *in vivo*. This, together with the expansion of computing technology during the 1980s, led to the first wave of structural neuroimaging studies aimed at extending previous postmortem results. Much of this early work utilized volumetric parcellation methods, whereby brain images are segmented according to different anatomical landmarks, and the volumes and tissue content (gray matter, white matter, cerebrospinal fluid) of these different regions are computed and compared between subject groups or throughout development. This parcellation step has been performed with a variety of methods, including the use of stereotactic coordinates (Jernigan *et al.*, 1991; Reiss *et al.*, 1996), manually drawn regions of interest (ROIs) (Giedd *et al.*, 1996c; Sowell *et al.*, 2002b), and automated protocols (Giedd *et al.*, 1999a, 1996a).

1.4.1 Gray matter decreases in development

Given the previous postmortem observations of regional and temporal variations in synaptic density and myelination throughout the brain, the gray and white matter volume estimates extracted through

these volumetric parcellation methods would be expected to show similar age-related developmental trajectories and regional differences. This was first demonstrated by Jernigan and Tallal, who observed that a group of children aged 8-10 had significantly more cortical gray matter than a group of young adults, as well as a higher gray matter to white matter ratio (Jernigan and Tallal, 1990). A subsequent study extended these findings to confirm that the group differences were due to continuous age-related decreases in gray matter volume with time — independent of brain size — and localized these effects to superior frontal and parietal cortices (Jernigan *et al.*, 1991). These studies marked the first *in vivo* morphological evidence in support of the earlier postmortem histological work by Huttenlocher and colleagues. While not a direct measure of synaptic density, the volumetric MRI finding of decreased gray matter volume is consistent with the regressive synaptic pruning changes previously described, and aligns with the theory that evolutionarily complex regions like the frontal lobe show more protracted timing in their development than evolutionarily simpler regions like primary motor/visual cortex. Even this early on, Jernigan and colleagues were also aware of the possible relationship between their *in vivo* MRI findings and the postmortem white matter myelination studies of Yakovlev and Lecours, and suggested that an "apparent" cortical thinning could be due, in part, to progressing myelination. Thus, a component of these observed changes might not be a gray matter loss, per se, but a transition of unmyelinated "gray" matter into white matter, which, on MRI, would appear as a gray matter volume "loss" during the childhood and adolescent years.

1.4.2 Regional and temporal dynamics

Since these initial observations of childhood and adolescent gray matter volume loss, other investigations have confirmed the general trend (Caviness *et al.*, 1996; Giedd *et al.*, 1999a; Ostby *et al.*, 2009; Pfefferbaum *et al.*, 1994; Reiss *et al.*, 1996; Sowell *et al.*, 2002b; Wilke *et al.*, 2007) and extended these observations in several important ways. In a large cross-sectional sample of 161 subjects aged 3 months to 70 years, Pfefferbaum and others were able to demonstrate the early rise and plateau in total brain volume by approximately age 5, as well as the late childhood and

adolescent decline in cortical gray matter volume. The extended age range of the sample allowed them to characterize the trajectory of the gray matter volume decline as curvilinear, which suggested an overall ∩ shaped curve consisting of early childhood gray matter increases followed by a relatively early peak, and then late childhood and adolescent reductions (Pfefferbaum *et al.*, 1994). This general time course of cortical development is a feature that has gone on to become one of the hallmarks of structural brain development (Courchesne *et al.*, 2000; Paus *et al.*, 2001; Sowell *et al.*, 2003).

Other studies have investigated the relative volume changes (controlling for global increases in total brain volume) more closely in broader age samples and in specific cortical and subcortical structures. In doing so, this work has demonstrated further heterogeneity in maturational timing and trajectory complexity across the brain. Importantly, the relative gray matter volume reduction during adolescence was confirmed to be most concentrated in the frontal and parietal lobes (Sowell *et al.*, 2002b). Meanwhile, subcortical gray matter structures like the basal ganglia also generally showed a relative volume reduction, although with a simpler linear trajectory than the cortex over the age range of late childhood to young adulthood (Ostby *et al.*, 2009; Sowell *et al.*, 2002b). Tazrouchi et al. recently applied some of the modern spatial normalization methods to align each individual's structural MRI brain data to a group average of all individuals studied, and conducted an analysis in the vein of these classical volumetric studies by examining volume change over time in over 100 regions throughout the cortex. Using an exponential nonlinear function to model the upstroke portion of the developmental curve, they were able to estimate the age at which these different gray matter regions reached full development. In doing so, they provided a compelling demonstration of the previously theorized maturational sequence, with primary somatosensory and visual cortices maturing the earliest, and then posterior-to-anterior and inferior-to-superior trends in the developmental timing of the remaining temporal, parietal, and frontal lobes (Tzarouchi *et al.*, 2009).

1.4.3 White matter increases in development

Interestingly, while gray matter volume was observed to peak early, researchers began to consistently observe that white matter volume continues to steadily increase roughly linearly from birth through adolescence and young adulthood (Caviness *et al.*, 1996; Paus *et al.*, 2001; Pfefferbaum *et al.*, 1994; Sowell *et al.*, 2002b; Wilke *et al.*, 2007). The timing of these changes shows a posterior-to-anterior gradient, which generally parallels the overlying gray matter, and has led to continued investigation into the interaction between these processes (Barkovich *et al.*, 1988). The white matter volume increase is consistent with the widespread reports of relative gray matter reductions during later childhood and adolescence, as a protracted increase in underlying white matter volume (due in part to increased oligodendroglial wrapping of axonal fibers) will increase total brain volume, and therefore decrease the relative gray matter volumes of specific structures compared to this total. The midsagittal corpus callosum was one of the first white matter areas to be examined in more detail, with volumetric analyses showing robust increases in total area throughout adolescence (Bellis *et al.*, 2001; Giedd *et al.*, 1999b) and a surprising anterior-to-posterior trend in the timing of the growth curve when the corpus callosum was subdivided into seven distinct segments (Giedd *et al.*, 1996a). This protracted nature of white matter development is a thread that we will see repeated in the following sections as imaging modalities and analysis techniques have advanced (Giedd *et al.*, 1999a; Lebel *et al.*, 2008b; Sowell *et al.*, 2003), and one that has gained increasing interest as more attention is being focused on the network properties of the brain as potential important mediators for the late cognitive development seen in domains like risk/reward processing, cognitive control, and working memory (Spear, 2000).

1.4.4 Sex differences

Sex-specific differences in brain structure were also extended with these structural imaging techniques. Total brain volume was confirmed to be approximately 10% larger in males than females at the plateau of overall brain volume that is reached during childhood (Caviness *et al.*, 1996; Courch-

esne *et al.*, 2000; Durston *et al.*, 2001; Gur *et al.*, 2002; Lenroot and Giedd, 2010), and the significant sex-specific effect remains even when height and weight are covaried (Giedd *et al.*, 1996b). However, more detailed regional volumetric observations of different cortical regions have been inconsistent — with varying reports of increased or decreased volumes in males and females that are further complicated by whether or not absolute or relative changes (to total brain volume differences) were reported (Sowell *et al.*, 2007).

Despite this variation, strong evidence suggests that there may be sexual dimorphism in the timing of the developmental trajectory in the cortex, such that males and females have similar overall trajectories of regional brain maturation (both with a $\cap$ shaped curve), but differing gender effects with time because of a difference in the timing of this trajectory (Giedd *et al.*, 1999a). Specifically, there appears to be approximately a 1–2 year phase difference between girls and boys, with peaks in gray matter occurring earlier in girls than boys, and regional variations in both phase and the actual differences between the sexes (Lenroot *et al.*, 2007). Because this is a temporally dynamic period of development (and not a static one, for example, like comparing fully-mature adults), assessing sex differences during childhood and adolescence has become a more complicated problem, which requires the dissociation of phase differences (particularly those caused by differences in age of pubertal onset) from sex differences in the maturational trajectories. An additional challenge is identifying those differences that persist into adulthood and actually have functional relevance.

Observations of sex differences in subcortical regions have also been somewhat more reproducible. In particular, over the course of development, the amygdala seems to increase in volume more in males, and the hippocampus more in females (Giedd *et al.*, 1997, 1996b,c; Wilke *et al.*, 2007). This is in line with animal studies that have shown high densities of steroid hormone receptors in the medial temporal lobe (Sarkey *et al.*, 2008), and also that sex steroids exert trophic effects on these structures (Cooke, 2006; Galea *et al.*, 2006; Zhang *et al.*, 2008). In one recent study specifically targeted to investigate the degree to which the rise of gonadal hormones during puberty contributes to the emergence of these sex-specific differences, we examined subcortical volume measures among a group of 80 adolescent boys and girls matched on sexual maturity within

a relatively narrow age range (Bramen *et al.*, 2011). This focused analysis revealed an interaction between sex and the effect of puberty in predicting amygdala and hippocampal volumes: While females actually had larger left amygdala and right hippocampal volumes than boys during early puberty (relative to total brain size), by late puberty the amygdala volume had increased in males but stayed relatively stable in females. In the right hippocampus, the effect of puberty was also to increase the volume in males, but surprisingly to decrease the volume in females. This reiterates the importance of considering the timing in the interpretation of developmental phenomena like sexual dimorphisms. Furthermore, these results suggest that the sex-specific differences in amygdala volume previously observed across a broader age sample (Giedd *et al.*, 1996c) are likely due in large part to the effects of puberty, but that the previous observation of relatively larger hippocampal volumes in females may be due to nonpubertal influences, as the direct contribution of puberty demonstrated in our recent observations would be to blunt this effect. These findings are particularly important in the context of adolescent brain development, as maturation of these processing centers, and their connections to areas like the prefrontal cortex, may contribute to the dramatic changes seen in social and emotional domains during this period of development (Dahl, 2004; Steinberg, 2005). The caudate nucleus has also been shown to be relatively larger, controlling for total brain volume, in females across several distinct samples (Giedd *et al.*, 1997, 1996b; Sowell *et al.*, 2002b; Wilke *et al.*, 2007). Put another way, the caudate is spared the reduction in volume that is typically shown by other structures in female brains. This observation of sexual dimorphism in the basal ganglia is also important, as it may relate to the emergence of similar sex differences in the incidence of several neuropsychiatric disorders (*e.g.* attention deficit hyperactivity disorder (ADHD), Tourette's syndrome) that are thought to involve these structures (Marsh *et al.*, 2008).

1.5 Brain mapping approaches

1.5.1 Advantages

The early volumetric MRI imaging observations by Jernigan and others helped lay the foundation for the next wave of neuroimaging studies designed to further characterize the anatomical changes that occur during normal development. While the volumetric protocols were able to validate much of the classical postmortem literature, as well as provide further evidence for gray matter loss, white matter gain, and regional/temporal dynamics, they are unable to precisely localize where these maturational changes are taking place within the relatively large regions of interest (ROIs) studied. Instead, these methods collapse entire regions of the brain down into one or several summary descriptive statistics that may not be characteristic of all functional and structural brain circuits within these large lobar regions. In contrast, newer methods like voxel based morphometry (VBM) and cortical thickness analysis are distinct in that they allow for statistical analysis at many points throughout the entire brain volume or at many points across the entire cortical surface, and the creation of whole-brain "maps" to visual these data. These enhanced analysis modalities, together with the traditional methods discussed above, have contributed greatly to our understanding of normal brain development and provide an important context for the further study of neurodevelopmental and psychiatric disease (Eliez and Reiss, 2000; Marsh *et al.*, 2008).

1.5.2 Voxel-based strategies

In voxel-based morphometry (VBM), the local fractional gray matter volume is analyzed in the neighborhood around each voxel in the brain to generate whole-brain maps of gray matter "density" or "concentration" (Ashburner and Friston, 2000). Spatial normalization algorithms are applied to align the brains of individual subjects so that each voxel then can be compared throughout development or between groups. Consistent with the previous volumetric studies and postmortem examples, whole-brain mapping strategies utilizing VBM show decreasing gray matter density during later development. In line with the coarse frontal and parietal lobar localizations of the earlier

volumetric reports, the regions showing the most protracted changes in these new analyses include clusters in the dorsal frontal and parietal cortices during the transition from childhood to adolescence (Sowell *et al.*, 1999a), as well as a distinct grouping of dorsal, medial, and orbital frontal cortical areas during the later transition from adolescence into young adulthood (See Figure 1.1) (Sowell *et al.*, 1999b). The relative specificity of these later changes to the frontal lobes is consistent with the similarly protracted time course of cognitive development in executive function domains, which are also typically thought to involve these frontal regions (Casey *et al.*, 2005; Luna *et al.*, 2004, 2010; Spear, 2000).

This notion of gray matter density was extended to allow for analysis on the cortical surface through the method of cortical pattern matching, where sulcal landmarks are manually identified and used to drive accurate nonlinear spatial normalization into a common template, while helping to account for regional, gender, and individual variability (Ashburner *et al.*, 2003; Luders *et al.*, 2004). Using this technique, protracted post-adolescent gray matter density decreases were again demonstrated in dorsal frontal cortex (Gogtay *et al.*, 2004; Sowell *et al.*, 2001b), and, for the first time, shown to correlate significantly with underlying relative brain growth in these regions (Sowell *et al.*, 2001b). This suggests the combined influences of both regressive processes like synaptic pruning, as well as progressive processes like myelination, are acting in these areas. Using similar gray matter density measurement techniques, and a powerful longitudinal design that tracked individuals prospectively for 8–10 years, Gogtay et al. provided further evidence that lower-order somatosensory and visual areas develop earlier than the higher-order association cortices that integrate these processes, and also that phylogenetically older areas develop earlier than younger areas (Gogtay *et al.*, 2004). Surprisingly, however, gray matter density increases were actually observed in bilateral perisylvian regions during the transition from adolescence to adulthood, and shown to correlate with both lateralized differences in sylvian fissure morphology and concomitant local brain growth (Sowell *et al.*, 2002a; Sowell *et al.*, 2001b). This suggests a particularly extended developmental trajectory in these gray matter regions beyond that in the dorsal frontal lobe, and perhaps implies a unique position for these canonical language areas in the developmental landscape

— with the typical inverse correlation between density and volume (decreasing density, increasing volume) reversed to give a direct relationship (increasing density, increasing volume) in these areas during this age range (Sowell *et al.*, 2003). Taken together, these findings again highlight both the regional as well as temporal complexities to the normal developmental sequence of brain structure.

1.5.3 Cortical thickness

The investigation of apparent cortical gray matter decreases during development reached full stride with the development of MRI cortical thickness measurement techniques. These automated algorithms extract mesh models of the white-matter/gray-matter boundary surface and the pial (*i.e.* cortical) surface, and then directly calculate the cortical thickness at many points throughout the cortical sheet (See Figure 1.2) (Fischl and Dale, 2000). Unlike the rather abstract interpretation of fractional gray matter "density" estimates, cortical thickness estimates are in physical units of millimeters and validate exceptionally well against the historical postmortem cortical thickness maps — with average measurements in children ranging from 1.5 mm in occipital cortex to 5.5 mm in dorsomedial frontal cortex (See Figure 1.2) (Sowell *et al.*, 2004a; Von Economo, 1929).

In addition to their strong agreement with postmortem data in terms of absolute thickness estimates, reports using this method are also in line with the mounting evidence from postmortem, volumetric, and VBM density measurements, which supports the picture that gray matter thickness peaks early and then declines due to a combination of progressive events like enhanced myelin penetration into the cortical neuropil, and regressive events like continued synaptic pruning (O'Donnell *et al.*, 2005; Shaw *et al.*, 2008; Sowell *et al.*, 2004a; Tamnes *et al.*, 2010). In a longitudinal study of 45 typically-developing children aged 5–11 years, who were scanned 2 years apart, these techniques were able to demonstrate gray matter thinning of 0.15–0.30 mm/year coupled to relative brain volume increases in right frontal and bilateral parieto-occipital regions (See Figure 1.3). This study was also able to reproduce the surprising earlier findings of gray matter increases in bilateral perisylvian language areas (Wernike's area), and extended these observations to the left inferior frontal gyrus — another language area (Broca's area) (Sowell *et al.*, 2004a). Cortical thickening was esti-

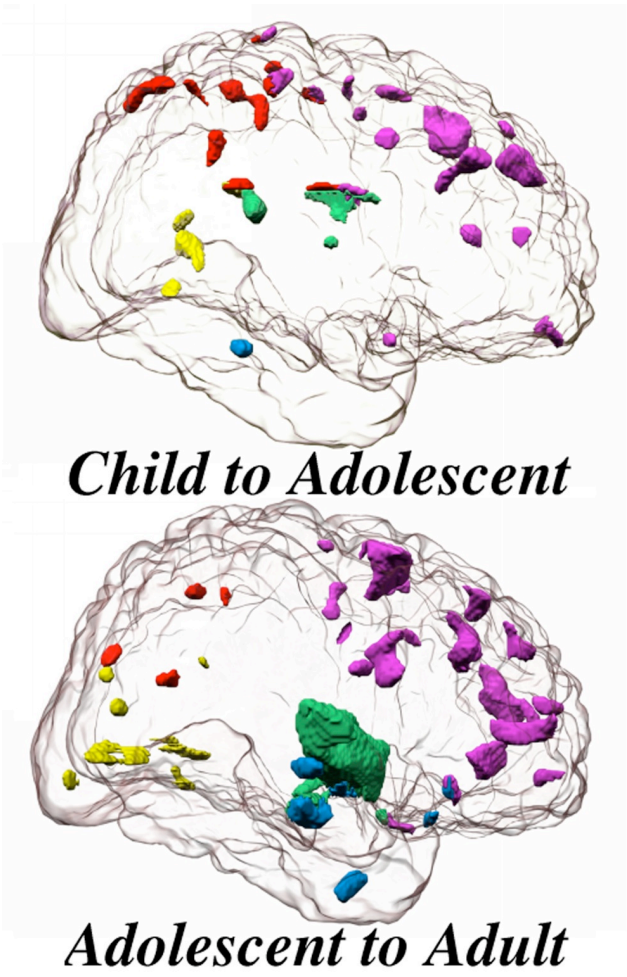


Figure 1.1: Gray matter density maturation. Voxel-based morphometry (VBM) measurements of fractional gray matter density/concentration show typical decreases during development. Colored volumes within a transparent cortical surface rendering represent the extent of significant decreases in gray matter density during the transition from childhood to adolescence (top panel) and adolescence to adulthood (bottom panel). Color-coding indicates which changes occurred in the frontal lobe (purple), parietal lobe (red), occipital lobe (yellow), temporal lobe (blue), and subcortical regions (green) (Sowell *et al.*, 1999a,b).

mated to be at a rate of 0.10–0.15 mm/year in these areas. This unique pattern of cortical thickening in the canonical language regions could be related to parallel gains in language processing made during this period of development. In another large longitudinal study of 375 children and young adults, changes in cortical thickness were modeled with a low-order polynomial basis set in order to investigate regional differences in the complexity of the developmental trajectory (Shaw *et al.*, 2008). Patterns of varying complexity were found to parallel the established histological maps of cytoarchitectonic complexity, and agree with the previous literature (Gogtay *et al.*, 2004; Sowell *et al.*, 2004a) — with simpler laminar areas like the limbic cortex having simpler trajectories, and more complex laminar areas like association cortex having more complex trajectories. Although cortical thinning during adolescence reflects developmental processes like myelination and synaptic pruning, it is important to note that cortical thinning also continues in some form throughout the rest of the lifespan (Sowell *et al.*, 2003). This likely belies a shift in etiology to degenerative changes associated with aging (Sowell *et al.*, 2004b), and recent work has sought to delineate this inflection point more precisely. By analyzing local gray and white matter signal intensities in the context of cortical thinning, the timing of the developmental peak was found to range from 8 to 30 years of age in different regions of the cortex, with the regional pattern following the general posterior-to-anterior gradient discussed before (Westlye *et al.*, 2010).

1.5.4 White matter

Even before the widespread adoption of diffusion imaging, which will be discussed in the next section, researchers were able to adapt traditional anatomical MRI analysis techniques to study white matter development (Wozniak and Lim, 2006). Magnetization transfer ratio (MTR) imaging is sensitive to the "bound" protons found on the phospholipids of myelin (Wolff and Balaban, 1989), and reflects the increasing myelination during early development (Engelbrecht *et al.*, 1998) as well as the posterior-to-anterior trend in the timing of this process (Buchem *et al.*, 2001). T2 relaxometry, which estimates the fraction of water in the brain that is associated with the phospholipid bilayer of myelin (MacKay *et al.*, 1994), has also been used to demonstrate the caudal-to-rostral wave of

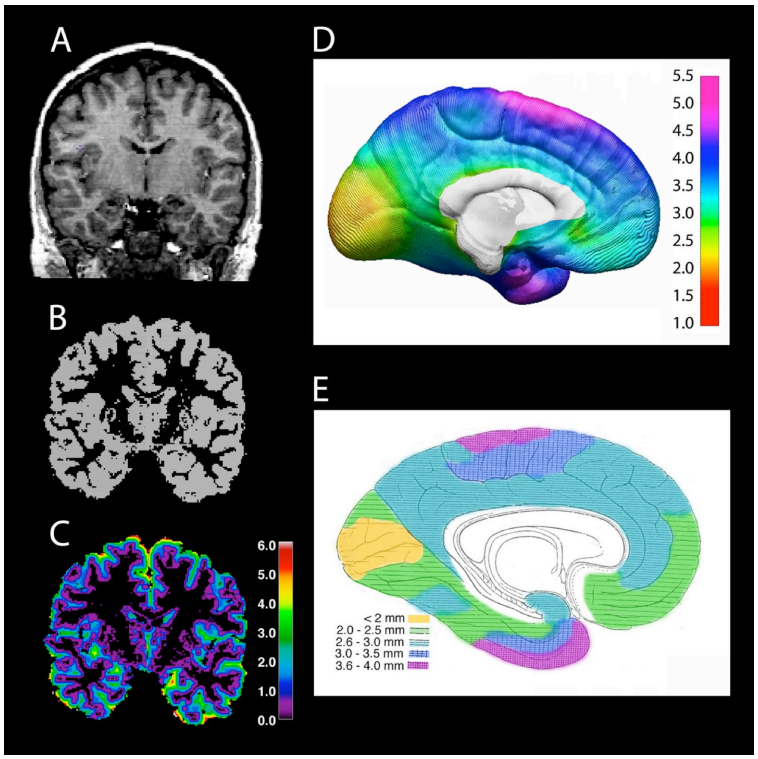


Figure 1.2: Cortical thickness analysis. Panels A-C show a single representative slice for one subject: (A) Raw T1-weighted anatomical MRI scan. (B) Gray/white matter tissue segmentation. (C) Gray matter thickness image, with thickness (mm) coded by color (warmer colors overlie the areas with the thickest cortex). (D) shows an in vivo average cortical thickness map generated by performing this analysis on a cross-sectional sample of 45 subjects. The brain surface rendering is color-coded according to the underlying cortical thickness (mm) and the color bar at right. The regional variations in cortical thickness can be compared to an adapted version of the classical Von Economo postmortem cortical thickness map (E), which has been color-coded in a similar manner over the original stippling pattern to highlight the similarity between the two maps (Sowell *et al.*, 2004a; Von Economo, 1929).

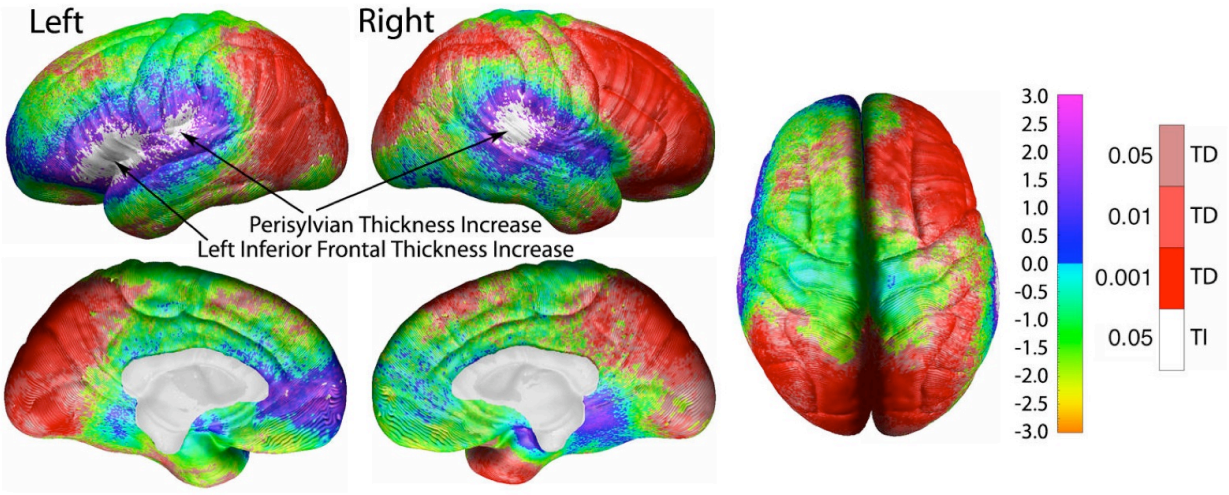


Figure 1.3: Gray matter thickness maturation. Statistical maps showing the significance of cortical thickness change in a longitudinal sample of 45 children scanned twice between the ages of 5 and 11. Areas showing significant thickness decrease (TD) are displayed in red, and areas showing significant thickness increase (TI) are displayed in white (See color bar and significance thresholds at right). Non-significant areas are coded by their t-statistic according to the left rainbow color bar. Arrows highlight the relative specificity of thickness increases during this age range to canonical language areas in the left inferior frontal gyrus (Broca's area) and perisylvian region (Wernike's area) (Sowell *et al.*, 2004a).

myelination (Lancaster *et al.*, 2003). In an application of the VBM technology to the white matter, Paus and colleagues were able to powerfully interrogate the rather general "global white matter increases" observation previously described to obtain a much richer localization of the precise anatomical regions involved. In an 88 subject sample of children aged 4–17 years, they observed a prominent increase in white matter density in the internal capsule bilaterally, as well as the left arcuate fasciculus, suggesting continued maturation of corticospinal and frontotemporal fibers through this age range (Paus *et al.*, 1999). This work agrees with the postmortem data from Yakovlev and Lecours, and demonstrates the unique progressive changes that are occurring in the white matter while the cortex has shifted to undergo predominantly regressive events. Confirming the surprising corpus callosum results of the classical volumetric study by Giedd et al. that was discussed earlier, Thompson and colleagues applied a continuum mechanics approach to obtain maps of local tissue deformation in the corpus callosum during development. Their longitudinal design studied 6 children aged 3–11 with a follow-up interval of up to 4 years, and again demonstrated an anterior-to-posterior wave in the timing of maximal local growth (See Figure 1.4) (Thompson *et al.*, 2000).

This contrasts with the general posterior-to-anterior trend that has been observed in gray matter cortical regions, and suggests a unique pattern of development in this region of interhemispheric fiber connectivity.

1.5.5 Sex differences

Continuing the trend from volumetric results, VBM gray matter density and cortical thickness observations of sex-specific effects during development have also been variable (Wilke *et al.*, 2007). However, this topic remains a critical issue, as sex-specific differences in brain development are likely to contribute to the sexually dimorphic susceptibilities to a variety of psychiatric disorders — like schizophrenia and major depression — that emerge during adolescence (Durston *et al.*, 2001; Lenroot and Giedd, 2010). Returning to the issue of gender differences in development, Sowell et al. analyzed cortical thickness and local brain size (taken as the distance from the center of the brain) in a large sample of 176 healthy subjects aged 7–87 years (Sowell *et al.*, 2007). In line with previous studies, male brains were larger than females at all locations. Strikingly, however, absolute cortical thickness was greater in females in right inferior parietal and posterior temporal regions even without accounting for the smaller overall size of female brains. This finding was not significantly modulated by age, and was demonstrated even more robustly across broad right temporal and parietal regions when an age- and brain volume-matched subset of 18 males and 18 females was evaluated (See Figure 1.5). These findings suggest that there are both regionally- and sex-specific differences in cortical thickness that appear relatively early in childhood, and support earlier reports of selective relative increases in gray matter volumes in females (Allen *et al.*, 2003; Goldstein *et al.*, 2001; Gur *et al.*, 2002; Im *et al.*, 2006; Nopoulos *et al.*, 2000; Sowell *et al.*, 2002b). Although the corpus callosum is also a frequent target for brain mapping research into sex-specific effects on brain development, no consensus has been reached and the topic remains frequently debated (Giedd *et al.*, 2006).

Because of the overall smaller brain volume in females, it has also been proposed that there may be evolutionary pressure to develop other compensatory mechanisms. Through sulcal delineation

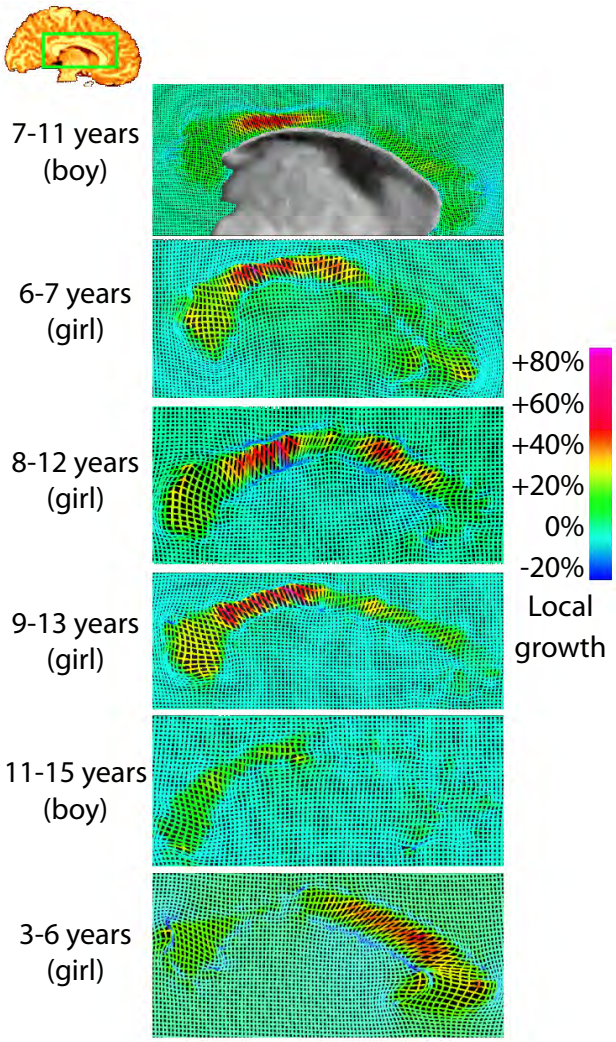


Figure 1.4: Corpus callosum maturation. Maps of the local volume changes in the corpus callosum are shown for six individuals aged 3–15 years, who were scanned twice longitudinally with an interval of up to four years. Maturation includes outward tissue expansion (warmer colors), with a dynamic wave in timing such that more frontal regions show prominent change early, and more posterior regions show prominent change later (Thompson *et al.*, 2000).

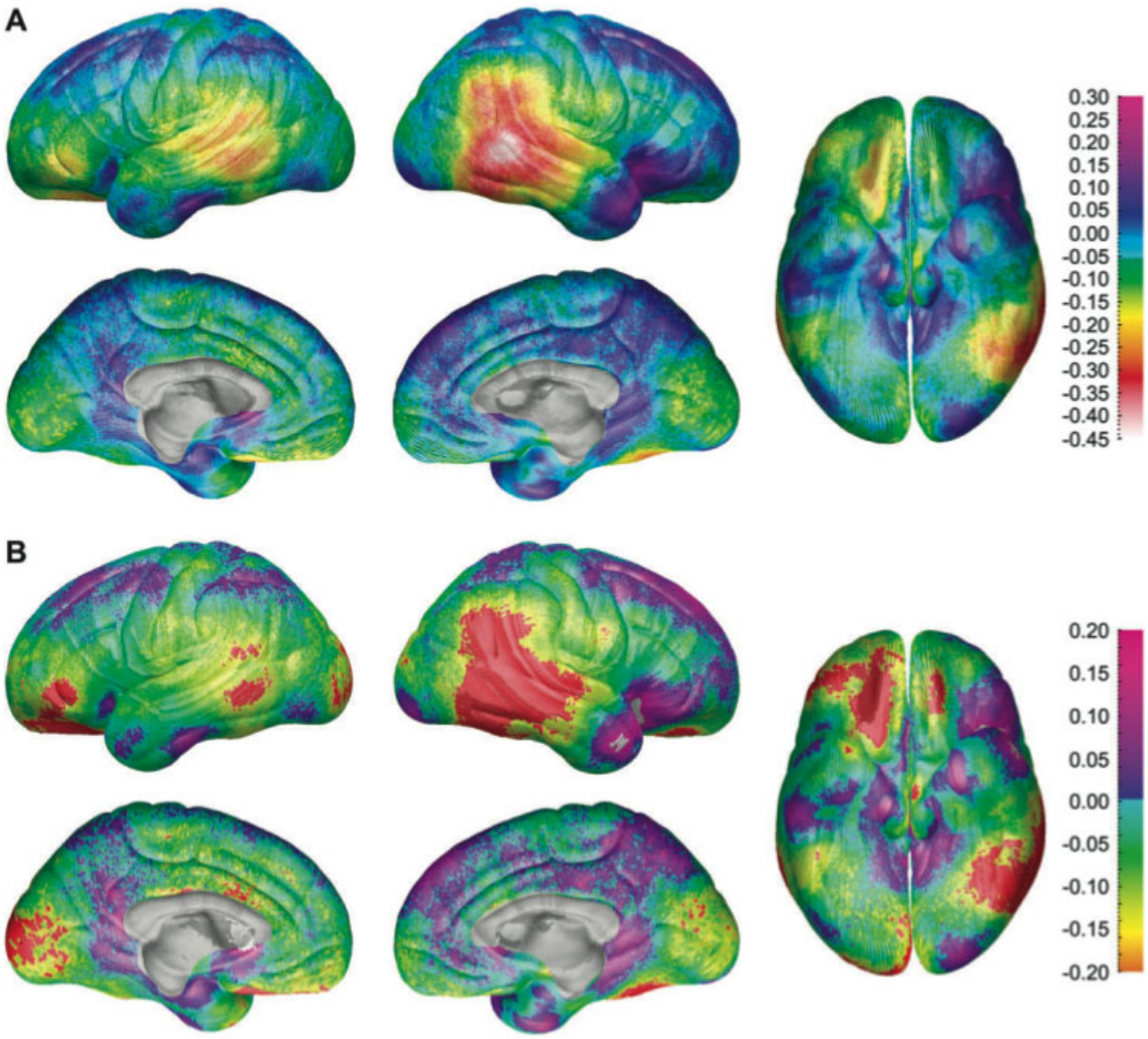


Figure 1.5: Sex-specific differences in cortical thickness. (A) Sex differences in cortical thickness (mm) among an age- and brain volume-matched sample of 18 males and 18 females. Warmer colors (<0 on the color bar at right) are regions where females have thicker cortex, and cooler colors (>0 on the color bar at right) are regions where females have thinner cortex, relative to males. (B) Statistical maps showing the significance of these sex differences. Areas where the cortex is significantly thicker in females are shown in red, and include right inferior parietal and posterior temporal, and left posterior temporal and ventral frontal regions. Areas where the cortex is significantly thinner in females are shown in white, and are limited to small regions in the right temporal pole and orbitofrontal cortex. The correlation coefficient is mapped for nonsignificant regions according to the color bar at right (Sowell *et al.*, 2007).

and cortical pattern matching techniques, it has been shown that females tend to develop a greater degree of cortical "complexity" by young adulthood (Luders *et al.*, 2004). This suggests that there is more cortical surface per unit volume in females, and may be one mechanism through which female brains have become optimized for their smaller size.

An increasing focus is also being shifted away from sex-specific differences, per se, to the known differences in pubertal timing and sex steroid levels that are likely to be major contributors to observed sex-specific effects and their frequently observed modulation by age (Giedd *et al.*, 2006; Lenroot *et al.*, 2007). The emerging picture suggests that puberty and sex steroids do indeed have organizing effects on brain development (Bramen *et al.*, 2011; Neufang *et al.*, 2009; Peper *et al.*, 2009a; Witte *et al.*, 2010). One recent study of 107 9-year-old monozygotic and dizygotic twin pairs noted strong overall heritability in regional brain volumes, but also demonstrated decreased frontal and parietal gray matter density among the subgroup of individuals who had begun to develop secondary sexual characteristics of puberty (Peper *et al.*, 2009b). Further investigation among the same cohort revealed that the serum level of luteinizing hormone, one of the first indicators of puberty, is associated with both increased overall white matter volume and increased white matter density in the cingulum, middle temporal gyrus and splenium of the corpus callosum (Peper *et al.*, 2008). The splenium observation is particularly intriguing, as this is the same region shown to have maximal growth over the 9–13 age range in a different study (Thompson *et al.*, 2000). These results — observed between otherwise very well-matched groups — suggest that the onset of puberty and sex steroid levels may directly contribute to the decreases in gray matter and increases in white matter that are prominent features of normal brain development during late childhood and adolescence.

1.5.6 Summary

Taken together, these structural imaging studies represent a powerful evolution in our understanding of brain development during childhood and adolescence. The overall picture remains one of early overall brain growth, followed by a transition around age 5 to gray matter decreases coupled to persistent white matter increases. These processes continue through adolescence, but relatively

balance each other in magnitude. Thus, while overall net brain volume changes relatively little past the age of 5, adolescence remains a period of dynamic change beneath the pial surface.

1.6 Diffusion MRI

One of the remarkable discoveries to emerge from these developmental neuroimaging studies is the continued expansion of white matter volume well into adulthood (Giedd *et al.*, 1999a; Sowell *et al.*, 2003). This robust and protracted increase has rewritten the age range associated with brain development (Pujol *et al.*, 1993), and has driven an increasing focus on the white matter and its network connectivity as a possible mediator for the late cognitive gains seen in executive function domains during typical development (Liston *et al.*, 2006), as well as a possible mechanism for neuropathology (Le Bihan, 2003) and training-induced increases in performance (Bengtsson *et al.*, 2005; Carreiras *et al.*, 2009).

1.6.1 Diffusion Tensor Imaging (DTI) theory

Simultaneously with this growing interest in studying the white matter, as it relates to connectivity between still-maturing brain regions and cognitive function, diffusion imaging was maturing as an MRI variation specifically tuned to examine the white matter (Basser *et al.*, 1994; Bihan *et al.*, 1986; Pierpaoli *et al.*, 1996). Since the diffusion properties of water within neural tissue are affected by the geometry of the neuronal microenvironment, it is intuitive that diffusion imaging can provide a sensitive lens through which the microstructural properties of the white matter can be investigated. Specifically, differences in microstructural properties like fiber coherence, axon packing, and myelination have all been shown to manifest as changes in the diffusion MRI signal (Beaulieu, 2002). By viewing this diffusion landscape within the brain from multiple angles, a more complete "tensor" model of diffusion can be generated for each voxel (Basser *et al.*, 1994). This can be thought of geometrically as a diffusion ellipsoid, with diffusion components in the radial (RD, radial diffusivity) and axial (AD, axial diffusivity) directions (See Figure 1.6). The size of this

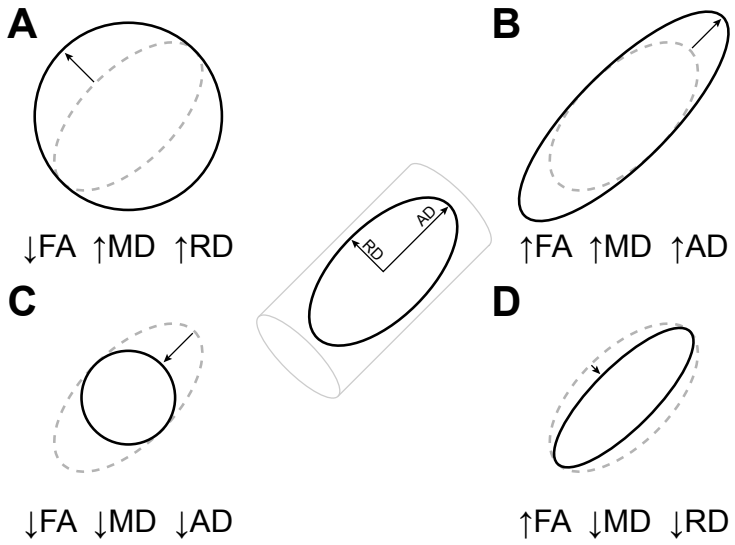


Figure 1.6: Diffusion Tensor Imaging (DTI) metrics. DTI metrics include fractional anisotropy (FA), which is a unitless measure of the directionality of diffusion, and mean diffusivity (MD), which is the overall magnitude of diffusion. The center panel shows a cross section of the DTI ellipsoid model of diffusion, which is assumed to be oriented along the fiber axis (shown here as a cylinder). Individual diffusion components along the axial (AD) and radial (RD) directions contribute to the FA and MD values at each point in the brain. Panels A-D show different changes in the individual diffusion parameters, and their varying effects on FA and MD. Note that changes in different diffusion components (AD or RD) can lead to the same effect on one diffusion metric, but have opposite effects on the other. Panel D represents the prevailing regime during development, where decreasing RD — due, in part, to advancing myelination — leads to increasing FA (a more pointed ellipsoid) and decreasing MD (a smaller ellipsoid).

ellipsoid corresponds to the overall mean diffusivity (MD). The shape of the ellipsoid corresponds to the directionality of diffusion, and is termed fractional anisotropy (FA). It can vary from 0, for perfectly isotropic diffusion, to 1, for perfectly anisotropic diffusion (*e.g.* the ventricles have low FA, while the corpus callosum has high FA). Because it has been shown to be sensitive to myelination, this FA metric has received considerable attention as a way to track the developmental maturation within the white matter and investigate disease. See Le Bihan (2003) for an excellent review.

1.6.2 Diffusion parameters in development

Using this unique framework, there has been a surge in research aimed at more deeply characterizing the normal developmental processes in these important regions of connectivity that were previously obscured by low contrast within the white matter on traditional T1-weighted anatomical

MRI. Similar to the general description in the overlying gray matter, the developmental trajectory within the white matter is both a nonlinear function of time, and has prominent regional variations (Lebel *et al.*, 2008b; Mukherjee *et al.*, 2001; Snook *et al.*, 2005). From birth, there is a rapid rise in diffusion directionality (FA; see Figure 1.7), coupled to a decrease in overall diffusivity (MD) (Bava *et al.*, 2010; Engelbrecht *et al.*, 2002; Hüppi *et al.*, 1998; Löbel *et al.*, 2009; Morriss *et al.*, 1999; Mukherjee *et al.*, 2001; Neil *et al.*, 1998; Schmithorst and Yuan, 2010; Schneider *et al.*, 2004). In an interesting contrast to this general pattern within the white matter, gray matter cortical regions actually have been observed to have decreasing FA in a sample of preterm infants (McKinstry *et al.*, 2002). This could reflect the fact that changes in FA are not highly specific for myelination, and may also occur in response to cortical maturational processes like synaptogenesis. Further, these observations may be related to the perinatal period of selective vulnerability in neural tissue, which has been demonstrated in animal studies and confirmed in humans through MRI (Miller and Ferriero, 2009). The white matter pattern of increasing FA and decreasing overall diffusion, although not universally reported in later development (Schneiderman *et al.*, 2007), generally continues in a decelerating fashion throughout childhood, adolescence, and in some areas, into adulthood (Bonekamp *et al.*, 2007; Klingberg *et al.*, 1999; Schmithorst *et al.*, 2002; Zhang *et al.*, 2007). There is a relatively stable plateau of these parameters during adulthood, and then eventual declines later in life (Davis *et al.*, 2009; Salat *et al.*, 2005). Accordingly, the developmental rising portion of this arc has been modeled as a linear (Snook *et al.*, 2005), polynomial (Hsu *et al.*, 2010), or exponential function (Lebel *et al.*, 2008b; Mukherjee *et al.*, 2001; Schneider *et al.*, 2004). The earliest reports utilized an ROI approach to look at diffusion properties averaged across specific anatomical locations, and were able to reproduce the "increasing FA, decreasing MD" pattern across a broad variety of regions within the brain and during different periods of development. In one example, Suzuki and colleagues examined ROIs placed bilaterally in the frontal and parietal white matter of 16 children and young adults. They observed increased FA and decreased overall diffusivity with age, but went on to make the important determination that the etiology of these changes in FA and MD was a primary decrease in both radial (RD) and axial (AD) diffusion com-

ponents, with a larger decrease along the radial direction (Suzuki *et al.*, 2003). This explains the overall decreased diffusivity that was observed (both components decreased), but also the increased diffusion directionality (one component decreased more than the other). The dominance of changes in radial diffusivity (RD) during development is an important phenomenon that has been broadly replicated (Giorgio *et al.*, 2008; Lebel *et al.*, 2008b; Löbel *et al.*, 2009; Qiu *et al.*, 2008), although not universally (Ashtari *et al.*, 2007; Giorgio *et al.*, 2010), and is thought to relate to the primary role that extended myelination plays during this age range (Song *et al.*, 2002).

Paralleling the advancements made in the analysis of the cortex, methods have quickly adapted to include whole-brain mapping techniques that are able to examine the brain in a spatially continuous manner and better localize developmental changes. In general, these later efforts using VBM and similar techniques have both confirmed and extended the earlier ROI findings of broadly increasing FA and decreasing MD (Snook *et al.*, 2007). Tract-based spatial statistics (TBSS) is an evolution of these methods that is tailored specifically to the analysis of DTI data, and has been used successfully to demonstrate age-related changes in diffusion imaging parameters (Bava *et al.*, 2010; Burzynska *et al.*, 2010; Giorgio *et al.*, 2010, 2008). By projecting the imaging data onto a tract "skeleton" consisting of the cores of the white matter tracts, TBSS avoids some of the alignment problems that arise when the high contrast FA maps are compared using traditional voxel-by-voxel techniques (Smith *et al.*, 2006b, 2007). In a sample of 75 children through young adults that were analyzed using this approach, widespread FA increases and diffusivity decreases were again demonstrated spanning the frontal, temporal, and parietal lobes, and the cerebellum (Qiu *et al.*, 2008). Recognizing the need to synthesize these reports into a normative reference standard against which to judge clinical abnormalities, effort has also been directed towards generating developmental brain atlases that integrate this diverse set of information (Hermoye *et al.*, 2006; Löbel *et al.*, 2009; Mori *et al.*, 2008; Verhoeven *et al.*, 2010).

1.6.3 Fiber tractography

By making the assumption that the direction of the diffusion ellipsoid (*i.e.* the direction of principle diffusion) is pointing in the same direction as the neuronal fiber axis, streamlines can be generated passing from voxel to voxel along the path of principle diffusion. In this manner, the DTI technology has been extended to allow for in vivo fiber tractography (Behrens *et al.*, 2003; Catani *et al.*, 2002; Conturo *et al.*, 1999; Mori *et al.*, 1999). This allows for individualized measurements to be made that are tailored to each subject's anatomy, which circumvents many of the problems associated with attempting to register a diverse set of brains to a single template. Although these algorithms have generally validated well against postmortem dissections for many major white matter tracts, specific limitations related to issues like partial volume averaging and complex fiber geometries must be considered (Pierpaoli *et al.*, 1996). Using this technology, together with standardized protocols for delineating the major white matter tracts of interest (Wakana *et al.*, 2007), researchers have mapped the development of white matter fiber connectivity from before birth (Huang, 2010; Huang *et al.*, 2009; Huang *et al.*, 2006), through childhood, adolescence, and adulthood (Behrens *et al.*, 2003; Liu *et al.*, 2010; Wakana *et al.*, 2004), and even through evolution (Rilling *et al.*, 2008). Like other developmental neuroimaging efforts, these data provide important insight into human brain development in their own right, and additionally serve as important normative markers against which pathology can be judged (Adams *et al.*, 2010; Lebel *et al.*, 2008a; Thomas *et al.*, 2009). In a seminal report on the typical developmental trajectories within 10 major white matter tracts in a large sample of 202 subjects aged 5 to 30 years, Lebel et al. observed continually increasing FA in all regions (generally approximated well by an exponential funtion), but regional variations in timing such that the time to reach 90% of the adult plateau varied from approximately 7 years old in the inferior longitudinal fasciculus to beyond 25 years old in the cingulum and uncinate fasciculus (See Figure 1.7) (Lebel *et al.*, 2008b). Overall, they note that frontotemporal connections were the slowest to develop. In a representative example of the degree of intersubject diversity that exists even within tracts, DTI tractography has been used to demonstrate lateralization of different white matter tracts (Bonekamp *et al.*, 2007). In one particular study, left lateralization was shown for the

arcuate fasciculus (temporoparietal part of the superior longitudinal fasciculus), with higher FA and more streamlines in the left hemisphere (Lebel and Beaulieu, 2009). These findings are in line with previous observations of left-lateralization of perisylvian regions (Geschwind and Levitsky, 1968; Pujol *et al.*, 2002), and are thought to relate to the left hemisphere language dominance that exists in the majority of the population. Interestingly, this same pattern has been demonstrated even in neonates, suggesting that the structural basis of left hemisphere language dominance is present long before the development of speech (Liu *et al.*, 2010). Previous morphometric findings of local volume increases within the corpus callosum (Giedd *et al.*, 1996a; Thompson *et al.*, 2000) have also been explored with tractography. In a large sample of 315 subjects aged 5–59 years, Lebel and others demonstrated the typical trajectory of increasing FA and decreasing MD in the fiber tracts leading from all midsagittal sections of the corpus callosum (Lebel *et al.*, 2010). They also observed an "outer-to-inner" trend in the timing of these maturational arcs, which contrasts with the anterior-to-posterior volumetric trend observed on T1-weighted MRI (Thompson *et al.*, 2000) and highlights the additional insight that can be uncovered when the full extent of a tract is considered.

1.6.4 Sex differences

Diffusion imaging also reveals sex-specific structural differences within the white matter (Lenroot and Giedd, 2010; Schmithorst *et al.*, 2008). In one tractography study of 114 children, adolescents, and young adults, Asato et al. found generally decreasing radial diffusivity (RD), and protracted maturation past adolescence, in projection and association fibers that included connections between the prefrontal cortex and the striatum. Furthermore, they observed that white matter microstructural maturation proceeded in parallel with pubertal changes, with females having overall earlier maturation of white matter tracts than males (Asato *et al.*, 2010). This suggests that there may be hormonal influences on white matter maturation, and that by considering these aspects, one may obtain a more appropriate estimate of developmental progress than by only considering chronological age. This notion is supported by concurrent findings with structural MRI that demonstrate white matter volume increases during adolescence, especially in boys, are affected by testosterone

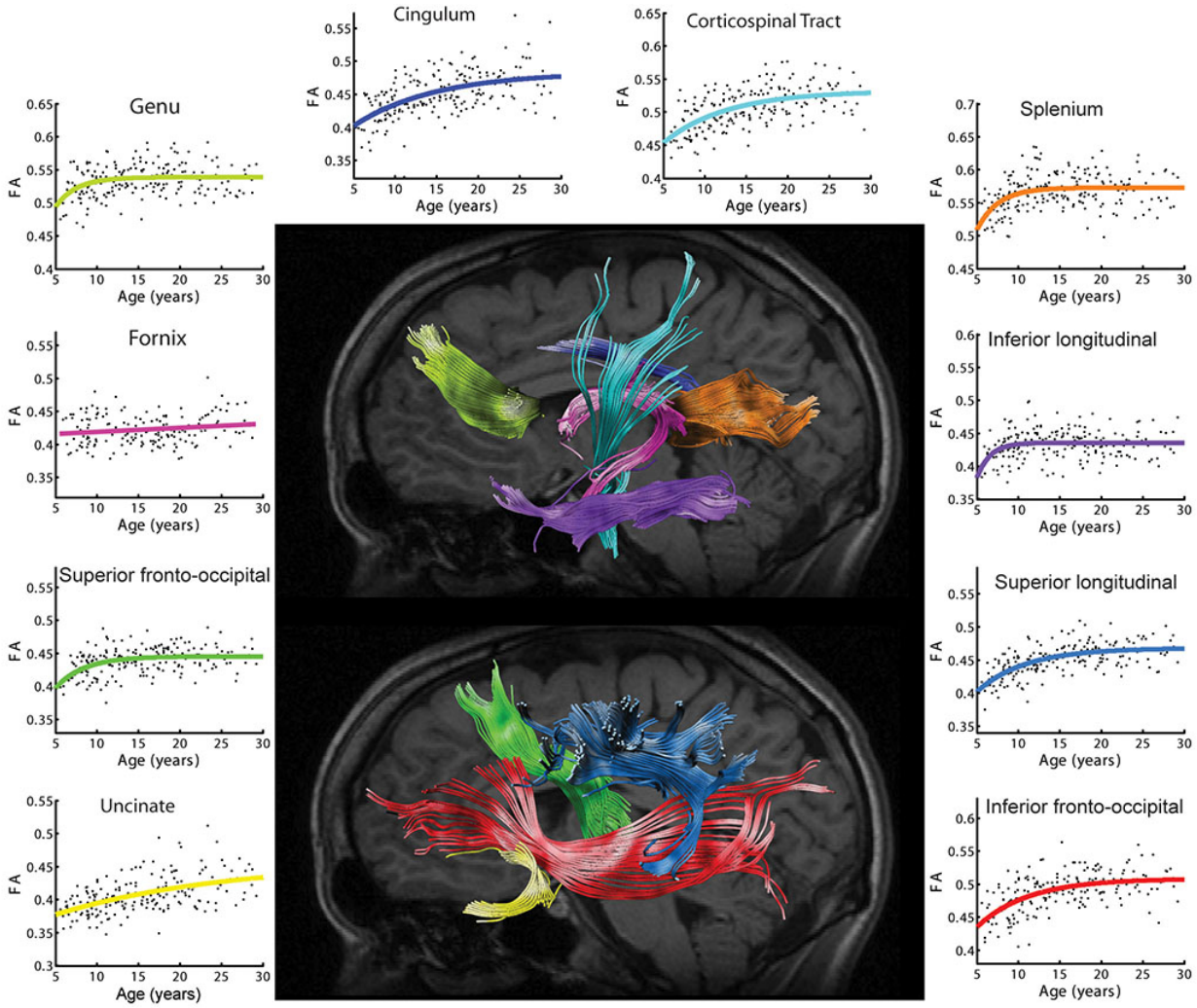


Figure 1.7: White matter maturation. Diffusion tensor imaging (DTI) tractography was used to identify 10 major white matter tracts in 202 individuals aged 5–30 years (center panels show the extracted tracts for a representative subject). Broad age-related increases in fractional anisotropy (FA), a DTI index of white matter maturation that is sensitive to myelination, were observed across all tracts. Maturational trajectories generally followed an exponential rise, with regional variations in mean FA as well as developmental timing. The surrounding scatterplots demonstrate these relationships, and are color-coded according to the tracts in the center panels (Lebel *et al.*, 2008b).

levels and androgen receptor genes (Paus *et al.*, 2010; Perrin *et al.*, 2008).

1.6.5 Summary

Taken together, diffusion imaging studies generally show increasing diffusion directionality (FA) and decreasing overall diffusion (MD) during development. These changes are predominantly due to decreasing diffusivity in the radial direction (RD) from the fiber axis, which suggests a primary role for myelination in this process. These changes progress rapidly from birth, through childhood, and eventually level off to a relatively stable adult plateau. Paralleling what has been observed in the cortex and through volumetric observations, there are regional variations in the timing of this developmental trajectory that follow a roughly posterior-to-anterior trend. Sexual dimorphism is also present, with females exhibiting earlier white matter maturation than males — a trend that mimics their differences in pubertal timing.

1.7 Connecting different techniques

1.7.1 Multimodal imaging

Although the development of cortical gray matter and the development of white matter microstructure have been investigated independently, one needs to consider their dynamics jointly in order to determine what relationships exist between them. This challenge returns to one of the original questions that stemmed from the postmortem histological findings — that is, "To what degree do myelination and synaptic pruning (and other cellular processes) contribute to the decreasing gray matter and increasing white matter that is found during brain development?" While these phenomena are undoubtedly linked, it remains unclear which is dominant and exactly how they interact. The maturation of DTI and structural MRI analysis techniques has now made it possible to investigate these questions using in vivo imaging data; however, in the end it will likely be necessary to complete the circle and validate these observations back in histological preparations.

In a study focusing on adolescence, Giorgio et al. began by using the TBSS method, discussed above, to demonstrate broad increases in FA that were driven predominantly by decreases in radial diffusivity (RD). They then made an important and innovative step by incorporating both DTI tractography and gray matter VBM to show that the putative fibers leading from the white matter regions with the strongest developmental effects connect with regions showing significantly decreased gray matter density in the cortex. Further, they observed that the gray matter density decreases were significantly correlated with the FA increases in the connected white matter (Giorgio *et al.*, 2008). By following the structural connectivity present in the actual data, and using these patterns to guide their comparisons, this protocol links the concurrent phenomena of white matter FA increases and gray matter density decreases more convincingly than was possible with previous qualitative visual inspections. Tamnes et al. investigated this same general question in a different manner by integrating cortical thickness, volumetric, and DTI measurements derived from a single sample of 168 participants aged 8–30 years (See Figure 1.8) (Tamnes *et al.*, 2010). As expected, they were able to demonstrate a combination of the phenomena seen in earlier individual studies, including broad cortical thickness decreases, white matter volume increases, FA increases (predominantly decreases in radial diffusion), and MD decreases. Most importantly, however, they were able to go on to demonstrate that, of the three measures, cortical thickness had the strongest relationship with age. Further, although the DTI and volume measures explained some of the variance in cortical thickness and each other, none of the measures were redundant. This implies that each may be sensitive to different microstructural processes, and that all are useful indicators of brain development and microstructural integrity (Fjell *et al.*, 2008). This reiterates the likely mixed regime of both synaptic pruning within the cortex, and advancing myelination at the gray-white cortical interface, which is contributing to brain morphological changes seen during adolescence. In another example, Choi et al. examined a completely different topic — general intelligence — but were able to gain similar benefits by integrating multiple imaging modalities. They observed that intelligence was generally related to cortical thickness and functional MRI (fMRI) blood flow response during a reasoning task. Because both sets of scans were performed on the same sample

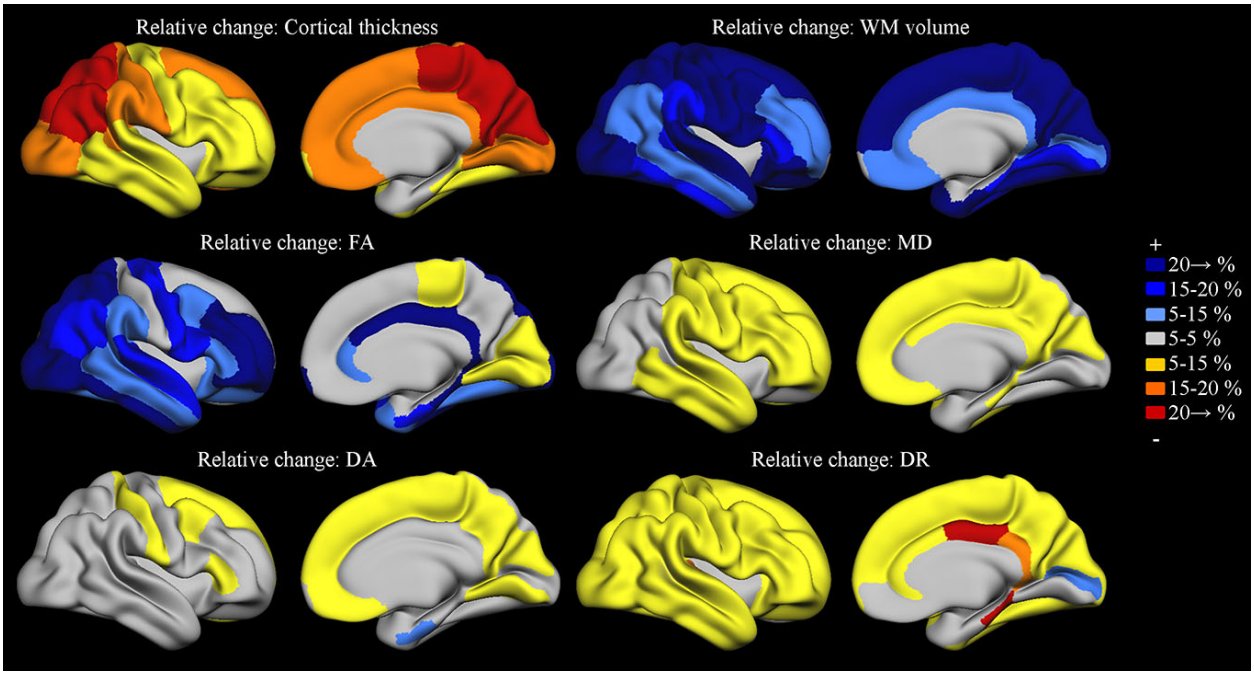


Figure 1.8: Multimodal Imaging: Volumes, cortical thickness, and DTI. Concurrent volumetric, cortical thickness, and diffusion tensor imaging (DTI) analyses were performed in the same sample of 168 participants aged 8–30. The percent changes in cortical thickness, white matter volume, fractional anisotropy (FA), mean diffusivity (MD), axial diffusion component (DA), and radial diffusion component (DR) are mapped by region and color-coded according to the color bar at right. Medial structures and corpus callosum are masked out (Tamnes *et al.*, 2010).

of subjects, however, the authors were able to go a step further — beyond this simple generalization — to observe that the crystallized component of intelligence was more strongly related to cortical thickness, while the fluid component of intelligence was more strongly related to functional blood flow response (Choi *et al.*, 2008).

1.7.2 Brain–behavior relationships

While important neuroanatomical insight can be gleaned from these structural brain mapping observations, perhaps the most significant outgrowth of this research has been an expanded understanding of the cognitive and behavioral changes that accompany this underlying maturation of brain structure. There has been a long tradition of investigation into the cognitive correlates of brain structure, but unfortunately many of the early findings — which commonly focused on differences between ethnic or social groups — are unreliable because of data collection and analysis bias

(Gould, 1978; Gould, 1981). With the advent of MRI, however, volumetric measurements of total brain size have shown a modest but reproducible correlation with general intelligence that emerges over the course of development (Peters *et al.*, 1998; Reiss *et al.*, 1996; Willerman *et al.*, 1991; Witelson *et al.*, 2006). However, the correlational nature of these findings does not at all suggest that groups with different brain sizes, like males and females, will have different intelligence. Indeed, independent of the possible relationships with neuroanatomy, it remains exceptionally controversial whether there is even any overall gender effect on intelligence (Blinkhorn, 2005; Hedges and Nowell, 1995; Irwing and Lynn, 2006; Jorm *et al.*, 2004; Lynn and Irwing, 2004; Neisser *et al.*, 1996), and if so, whether the small effect magnitudes that have been reported are relevant given the possible biases that may have contributed. An important additional phenomenon to consider is that both brain structure and intelligence are highly heritable (Shaw, 2007; Thompson *et al.*, 2001). Both are further impacted by environmental influences in a process that begins in utero, continues throughout life, and contributes to individual variations in structural brain development and cognitive function that exist even among monozygotic twins. Although not exclusive, the orchestration of structural brain development by these genetic and environmental factors is one way in which they can converge to influence cognitive development (Toga and Thompson, 2005).

Since there is evidence that brain development takes place through selective elimination and connectivity optimization, with prominent regional and temporal variability, it is not surprising that a global measure like total brain volume may not be the optimal choice for investigating the structural basis of cognitive development. Fortunately, the brain mapping strategies, discussed above, have had more success examining brain-region-specific relationships between structure and function. This work has supported many of the classical structure-function relationships discovered through lesion studies — for example, that the prefrontal cortex is related to cognitive control (Damasio *et al.*, 1994) — and also has extended these findings by 1) providing more detail, 2) including more normative subjects without pathology, and 3) allowing for broader investigation in the pediatric population. In this way, these modern neuroanatomical imaging studies, together with complementary results from functional neuroimaging (fMRI) methods that can measure task-

dependent blood flow response within the brain (Casey *et al.*, 1995; Luna *et al.*, 2010), have formed a powerful framework to investigate how brain development relates to cognitive function during childhood and adolescence. In this vein, continued investigation into the structural basis of general intelligence has revealed age-variable relationships between IQ and regional brain structure. In line with the total brain volume results, a correlation between IQ and gray matter volume develops by adulthood (Wilke *et al.*, 2003). However, regional relationships between IQ and gray matter structural measures appear earlier, and have been reported to include the anterior cingulate during childhood (Wilke *et al.*, 2003), the orbitofrontal cortex during adolescence (Frangou *et al.*, 2004), and the frontal lobe — particularly the prefrontal cortex — by adulthood (Haier *et al.*, 2004; Reiss *et al.*, 1996; Thompson *et al.*, 2001). Interestingly, these regional relationships between gray matter development and IQ appear to be modulated by sex, although the specific regions reported to be most associated with IQ for each sex have been variable (Haier *et al.*, 2005; Narr *et al.*, 2007). In one important study, which investigated the relationship between cortical thickness maturation and IQ in a large longitudinal sample of 307 children and adolescents, IQ was observed to correlate most closely not with cortical thickness, per se, but rather with the shape of the developmental trajectory in cortical thickness change (See Figure 1.9) (Shaw *et al.*, 2006a). The subjects that had the highest IQs tended to have the most dynamic cortical maturation, with more rapid cortical thickening during early childhood, and more rapid cortical thinning during late childhood and adolescence. However, in terms of absolute thickness, the superior intelligence group actually had thinner cortex at the start of the age range studied (approximately age 7), peaked later, and then had relatively equal thickness to the others by the end of the age range (approximately age 19). This observation highlights the notion that, like the pattern of structural maturation itself, the relationships between brain structure and cognitive ability are complicated by their dependency on age during the course of development. While the specific pattern and methodologies of these studies have varied widely, the common pattern that has emerged is a relationship between frontal lobe structural brain development and general intellectual ability.

Other studies have investigated more specific cognitive functions and their relation to gray mat-

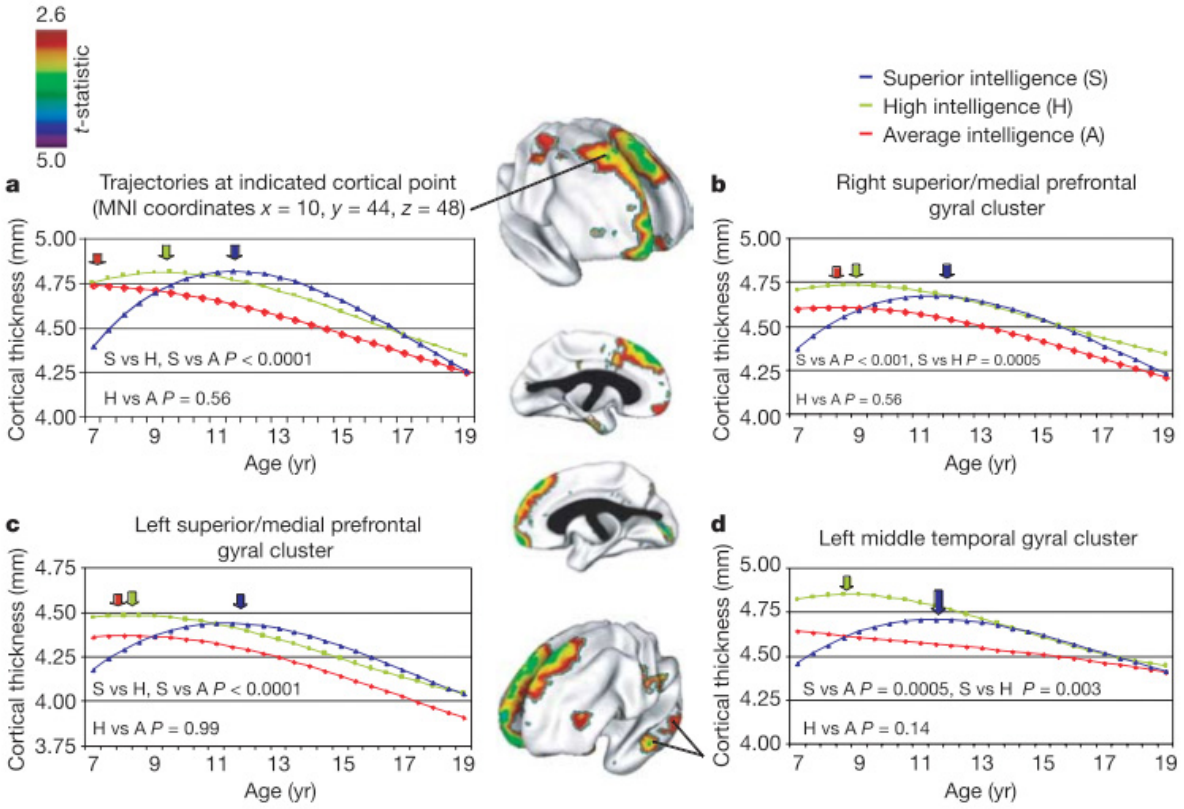


Figure 1.9: Trajectory of cortical thickness change vs. IQ. Higher IQ was associated with a more dynamic trajectory (more rapid thickening and thinning) in cortical thickness maturation among a sample of 307 children and adolescents scanned longitudinally. The center panel shows regions where there was a significant interaction between IQ group (superior, high, or average) and a cubic age^3 term in the regression analysis, which implies a varying trajectory shape in these regions. These individual trajectories are plotted in panels a–d, and are color-coded according to intelligence group. Arrows indicate the age at peak cortical thickness for each trajectory (Shaw *et al.*, 2006a).

ter structure. In the same longitudinal sample of 45 typically-developing children that was described previously, we observed inverse correlations between performance on the vocabulary subtest of the Weschler Intelligence Scale for Children (Wechsler, 2003) — a test of general verbal intellectual functioning — and gray matter thickness in left dorsolateral frontal and lateral parietal regions (See Figure 1.10) (Sowell *et al.*, 2004a). This is consistent with the language dominance of the left hemisphere, and suggests a possible relationship between these concurrent structural and cognitive developmental processes. While originally interpreted as possibly relating to developmental cortical thinning, the results of the Shaw *et al.* (2006a) study suggest that the individuals with the greatest verbal intellectual function here may still have been on the upstroke of their developmental arc in our much younger sample (age 5–11 years), and simply had thinner cortex at the time sampled. This nuance is also reflected in another study, which had an older sample (age 6–18 years) during the later period of development where increased cortical thickness is associated with higher IQ (Karama *et al.*, 2009). Further studies, again in the young sample of 5–11-year-olds, have investigated even more targeted cognitive subtests, including phonological processing, and motor speed and dexterity (Lu *et al.*, 2007). Structural development in the inferior frontal gyrus (a phylogenetically more complex area that matures slower and is still on the upward stroke of cortical thickening) was expected to relate to advances in phonological processing, which has been shown to involve this area on functional imaging studies (Bookheimer, 2002), but not to relate to advances in motor processing. Conversely, structural development in the hand motor region (a phylogenetically simpler area that matures earlier and is already experiencing cortical thinning) was expected to relate to advances in motor processing but not phonological processing. This predicted double dissociation was demonstrated as expected, which not only illustrates a specific alignment between language development and structural development in the inferior frontal gyrus, but also reiterates the regionally specific definition of "structural development" during childhood — with some cortical regions thinning, but some relatively specific language areas still exhibiting thickening. A similar analysis has also revealed relationships between cortical thinning and both delayed verbal recall functioning and visuospatial memory ability, which is again consistent with the functional

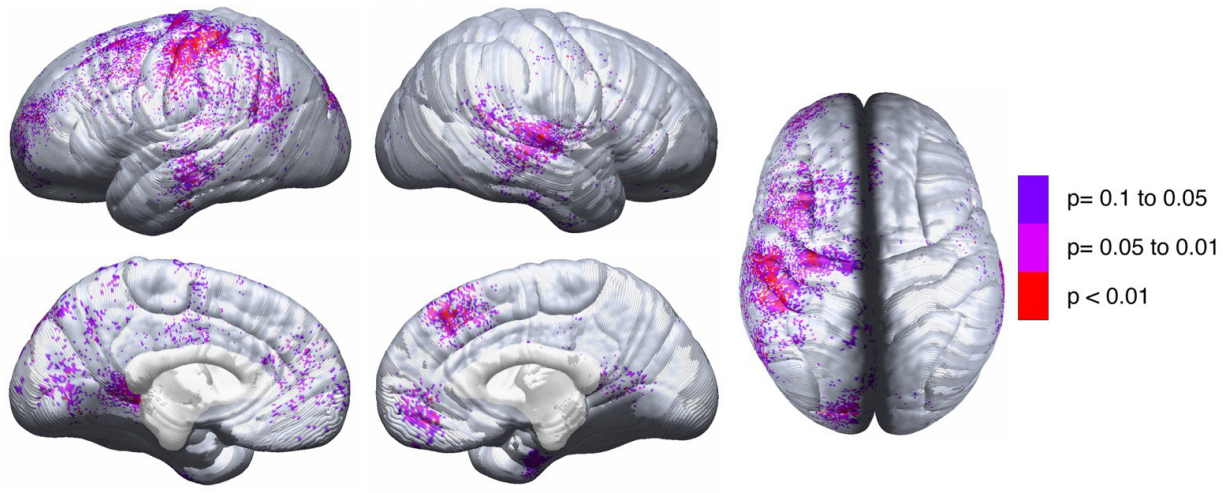


Figure 1.10: Cortical thickness vs. language functioning. Statistical maps showing the significance of the relationship between changes in cortical thickness and changes in vocabulary scores in a longitudinal sample of 45 children scanned twice between the ages of 5 and 11. Areas with a significant negative relationship (cortical thinning was associated with improved language performance) are color-coded according to their P value, with the significance thresholds shown in the color bar at right. No positive correlations were observed (Sowell *et al.*, 2004a).

neuroimaging literature that suggests the dorsolateral prefrontal cortex is involved with memory recall (Casey *et al.*, 1995; Sowell *et al.*, 2001a). The relationship between cognitive development and structural brain development is further supported by intervention/training studies, which suggest that even relatively short periods of cognitive or motor training can be associated with at least short term morphological changes in brain structure (Draganski *et al.*, 2004).

Diffusion imaging indicators of white matter development also relate to cognitive function. In a sample of 23 children and adolescents, there was a significant direct relationship between diffusion characteristics (FA) and working memory ability in inferior frontal and temporo-occipital regions, and the genu of the corpus callosum (Nagy *et al.*, 2004). This relationship existed above and beyond the correlation of each individual measure with age, which suggests that the maturation of the white matter in specific areas — as indexed by FA — may play a role in the development of (or simply reflect the development of) specific cognitive domains. In a similar design, others have shown correlations between Chinese reading score and FA in the anterior limb of the left internal capsule, and English reading score and FA in the corona radiata (Qiu *et al.*, 2008). In the

arcuate fasciculus lateralization tractography study discussed earlier, greater leftward lateralization was associated with better performance on cognitive tests of receptive vocabulary and phonological processing (Lebel and Beaulieu, 2009). These studies suggest that diffusion imaging is not only a useful technique for tracking normal anatomical maturation within the white matter, but also that regional DTI metrics can provide reflections of cognitive development in specific domains.

1.8 Conclusions and future directions

Our understanding of human brain development has accelerated over the last 20 years through the use of MRI and in vivo human brain mapping. Postmortem and histological studies have demonstrated that brain maturation, on the cellular level, encompasses both progressive and regressive events. These include synaptic pruning and protracted myelination, which continue to shape the underlying neural microstructure and regional brain morphology long after overall brain volume begins to plateau around age 5. Brain development, in general, can be characterized as both non-linear with respect to time, and also variable with respect to region. The hallmark of structural brain development during childhood is a striking change in the relative proportions of gray and white matter — with a peak and then decline in gray matter volume and cortical thickness, but a relatively sustained increase in white matter beyond adolescence. Across these different regions, there is a general posterior-to-anterior and inferior-to-superior trend in the timing of maturation, such that primary somatosensory and phylogenetically older areas of the brain tend to mature earlier than higher-order association cortices — particularly areas in the frontal lobe. Within the white matter, diffusion imaging indicators show decreasing diffusivity (MD) and increasing directionality (FA), which suggests that myelination continues through young adulthood and perhaps even beyond. Performance across a variety of cognitive domains has also been shown to relate to these structural changes, with the specificity of these relationships generally in line with classic functional neuroanatomical localizations.

Although the complexity of the regional and temporal patterns of structural brain development

makes investigating and interpreting these brain-behavior relationships challenging, future work should continue to focus on the possible functional manifestations of structural brain development. Particularly, by integrating different structural and functional imaging modalities with thorough cognitive assessments, we can investigate the ways in which these processes interact with each other within a more inclusive framework that more realistically encompasses the full developmental landscape. With the increasingly broad array of radiological features of development that have been characterized, there is additionally a growing need to reintegrate a firm neurobiological understanding of the cellular mechanisms that facilitate these changes. Finally, effort should continue to be directed towards uncovering the ways in which this basic neuroscientific knowledge concerning human brain development can be translated into a better context for the understanding and clinical treatment of neurodevelopmental disorders.

CHAPTER 2

Prenatal methamphetamine exposure

2.1 Abstract

Little is known about the effects of prenatal methamphetamine exposure on white matter microstructure, and the impact of concomitant alcohol exposure. Diffusion tensor imaging and neurocognitive testing were performed on 21 children with prenatal methamphetamine exposure (age 9.8 $\pm$ 1.8 years; 17 also exposed to alcohol), 19 children with prenatal alcohol but not methamphetamine exposure (age 10.8 $\pm$ 23 years), and 27 typically-developing children (age 10.3 $\pm$ 3.3 years). Whole-brain maps of fractional anisotropy (FA) were evaluated using tract-based spatial statistics. Relative to unexposed controls, children with prenatal methamphetamine exposure demonstrated higher FA mainly in left-sided regions, including the left anterior corona radiata (LCR) and corticospinal tract ($P < 0.05$, corrected). Relative to the alcohol-exposed group, children with prenatal methamphetamine exposure showed higher FA in frontotemporal regions — particularly the right external capsule ($P < 0.05$, corrected). Post-hoc analyses of these FA differences showed their etiology to be more about lower radial diffusivity (RD) than higher axial diffusivity (AD). We failed to find any group-performance interaction (on tests of executive functioning and visuomotor integration) in predicting FA; however, FA in the right external capsule was significantly associated with performance on a test of visuomotor integration across groups ($P < 0.05$). This report demonstrates unique diffusion abnormalities in children with prenatal methamphetamine/polydrug exposure that are distinct from those associated with alcohol exposure alone, and illustrates that these abnormalities in brain microstructure are persistent into childhood and adolescence — long after the teratogenic polydrug exposure in utero.

Key words: methamphetamine, white matter, alcohol, diffusion imaging, teratogen, in utero

2.2 Introduction

Methamphetamine (MA) abuse is a significant medical and social problem worldwide — with broadening use and manufacture in developing regions of Southeast Asia and Oceania (McKetin *et al.*, 2008), and continued prevalence in established centers like Japan, Taiwan, Hawaii, and the southwest mainland United States (Maxwell and Rutkowski, 2008; *SAMHSA* 2008). While methamphetamine use in adults has been clearly linked to broad negative effects on the central nervous system, as well as negative social outcomes and effects on other organ systems (McCann *et al.*, 1998; Thompson *et al.*, 2004), surprisingly little is known about the effects of prenatal exposure to MA on the developing brain (Thompson *et al.*, 2009) (See Roussotte *et al.* (2010), for a review of the available evidence). Recently, a large prospective study has demonstrated fetal growth restriction in the context of prenatal MA exposure, and has expanded observations of poorer neurobehavioral outcomes to include depressed arousal and movement scores, and higher stress in newborn infants (age 2.0 $\pm$ 1.6 days) (Lagasse *et al.*, 2011; Nguyen *et al.*, 2010; Smith *et al.*, 2008, 2006a). The first neuroimaging protocol that specifically addressed prenatal exposure to MA utilized [^{1}H]proton magnetic resonance spectroscopy (MRS) to demonstrate findings suggestive of metabolic abnormalities in the striata of exposed children (age 8.1$\pm$0.8 years) (Smith *et al.*, 2001). This was followed by a volumetric analysis using magnetic resonance imaging (MRI) that showed smaller subcortical volumes in the basal ganglia and hippocampi of affected children, and correlations between brain volumes and poorer performance on attention and verbal memory tests (age 6.9$\pm$3.5 years) (Chang *et al.*, 2004). Recent functional MRI (fMRI) evidence in children exposed to methamphetamine prenatally also suggests abnormal patterns of brain activation, including more diffuse brain activation during a verbal working memory task (age 9.5 $\pm$ 1.9 years) (Lu *et al.*, 2009), as well as lower frontostriatal activation during a visual working memory task (age 9.2 $\pm$ 1.8 years) (Roussotte *et al.*, 2011). The only published reports of white matter abnormalities include a region-of-interest dif-

fusion imaging study, which examined 3 and 4-year-old MA-exposed children, and showed lower diffusion in frontal and parietal areas, and a trend towards greater diffusion fractional anisotropy (FA) in the left frontal white matter of the exposed group (Cloak *et al.*, 2009). In an overlapping sample, ^{1}H-MRS also demonstrated higher metabolite concentrations (total creatinine, N-acetyl compounds, glutamate/glutamine) in frontal white matter (Chang *et al.*, 2009). Taken together, the disturbances in infant behavior and brain imaging data in children suggest that prenatal MA exposure negatively impacts brain development. However, conclusions about the specific effects of MA are limited, given that polydrug exposure is common in this population.

Concurrent prenatal alcohol exposure is particularly concerning because it is a known teratogen (Jones *et al.*, 1973), and has been shown to induce lasting clinical deficits (Spohr *et al.*, 2007). Furthermore, observations have shown that nearly half of the women who use MA during pregnancy also drink alcohol (Smith *et al.*, 2006a). Neuroimaging findings in children with heavy prenatal alcohol exposure include global, regional, and subcortical volumetric abnormalities, as well as cortical thickness and fMRI abnormalities (Coles and Li, 2011; Lebel *et al.*, 2011). Most relevant to our present report, a variety of white matter abnormalities have also been reported among children with heavy prenatal alcohol exposure, including, most prominently, gross deformities of the corpus callosum (Sowell *et al.*, 2010; Wozniak and Muetzel, 2011).

Here we studied the effects of prenatal methamphetamine exposure on white matter microstructure using whole-brain diffusion tensor imaging (DTI). By measuring the diffusion properties of water inside the brain, which are affected by constraints placed by the neuronal microenvironment, DTI is able to provide an indirect noninvasive characterization of white matter microarchitecture in vivo. Given the known impact of MA exposure on striatal structures in adult abusers, and limited evidence in children, we expected abnormalities in regions of white matter tracts that connect striatal with cortical structures. Given reports of deficits in executive measures of attention, deficits in visual motor integration (Chang *et al.*, 2004), and findings of higher diffusion anisotropy in frontal white matter in children with prenatal MA exposure (Cloak *et al.*, 2009), we expected a similar pattern in our older sample of children on tests of executive function (Trails B), visuomotor inte-

gration (VMI), and whole-brain FA maps. In spite of these hypothesized group differences, within groups we still expected regionally-specific relationships between FA and performance, such that higher FA (indicative of white matter fiber organization) in the frontal lobe would be associated with better performance (more efficient processing) on a test of executive function (Trails B), but not on a test of visuomotor integration (VMI), and, conversely, that higher FA in the parietal lobe would be associated with better performance on the VMI test, but not the Trails B test.

2.3 Methods

2.3.1 Participants

Participants were classified into three groups according to exposure status: 1) Methamphetamine-exposed subjects (MA, $n = 21$), 2) alcohol-exposed subjects (ALC, $n = 19$), and 3) typically-developing controls (CON, $n = 27$). Subjects were included in the MA group if they had exposure to methamphetamine based on parent/guardian report, or maternal or infant medical records. In line with previous literature on fetal alcohol spectrum disorders (FASDs) that recognizes the impact of frequent drinking, as well as less frequent but heavier drinking, participants were included in the ALC group if they had exposure to $\geq$4 drinks on any occasion or were exposed to $\geq$14 drinks in any week during the pregnancy (a "drink" is defined as a 12 oz. beer, 4 oz. glass of wine, or cocktail with 1 shot of liquor), and had no methamphetamine exposure (Hoyme *et al.*, 2005). Phone screening exclusion criteria for all groups included: 1) age younger than 5 years (we were most interested in the long-lasting effects of MA, and, additionally, available staff were only trained on this age range for neuropsychological testing); 2) IQ less than 70; 3) head injury with loss of consciousness over 20 minutes (although no milder head injuries were reported on a follow-up parent self-report questionnaire either); 4) physical (*e.g.* hemiparesis) or psychiatric illness, or developmental disability (*e.g.* autism) expected to prevent completion of the scanning or neuropsychological testing sessions; 5) other potential known causes of mental deficiency (*e.g.* chromosomal disorders); and 6) presence of implanted metal in the body. ALC subjects were largely recruited from a university-

associated social skills training group for children with FASDs. Subjects in the MA group were recruited from three sources: 1) Children of mothers who were in an MA rehabilitation program, 2) the same social skills group described above for the FASD subjects (after it was discovered that some of the mothers also had abused MA during pregnancy), and 3) self-referral in response to advertisements and word-of-mouth. CON subjects were recruited from the same Los Angeles communities as the exposed groups, and effort was made to recruit from similar socioeconomic status (SES) strata (*e.g.* advertisements targeted to zip codes with similar SES as our exposed subjects). Details of diagnostic procedures for fetal alcohol spectrum disorders used to classify ALC and MA subjects are described in another report (O'Connor *et al.*, 2006). Briefly, an experienced clinician examined alcohol-exposed children using the Diagnostic Guide for Fetal Alcohol Syndrome (FAS) and Related Conditions (Astley, 2004). This system uses a 4-digit diagnostic code reflecting the magnitude of expression of four key diagnostic features of FAS: 1) growth deficiency; 2) the FAS facial phenotype, including short palpebral fissures, flat philtrum, and thin upper lip; 3) central nervous system dysfunction; and 4) gestational alcohol exposure. Using these criteria, children with alcohol exposure (with or without concomitant MA exposure) were diagnosed with fetal alcohol syndrome (FAS), partial FAS, sentinel features, or alcohol-related neurodevelopmental disorder (ARND) (Figure 2.1). Following a complete description of the study protocol, parent/guardian consent and participant assent were obtained in accordance with procedures approved by the Institutional Review Board at UCLA.

2.3.2 Neuropsychological testing

Subjects underwent a broad neuropsychological testing battery administered by trained fulltime staff that were blinded to subject exposure status. Included among the tests were measures of general intelligence (prorated full-scale IQ) (Wechsler, 2003), visuomotor integration (VMI), and executive control (Trail Making Test). The VMI test instructs subjects to draw a series of geometric figures that are presented visually, and thus, performance reflects intact visual sensory input, motor output, and their integration (Beery, 1997). The Trail Making Test is a popular compound measure

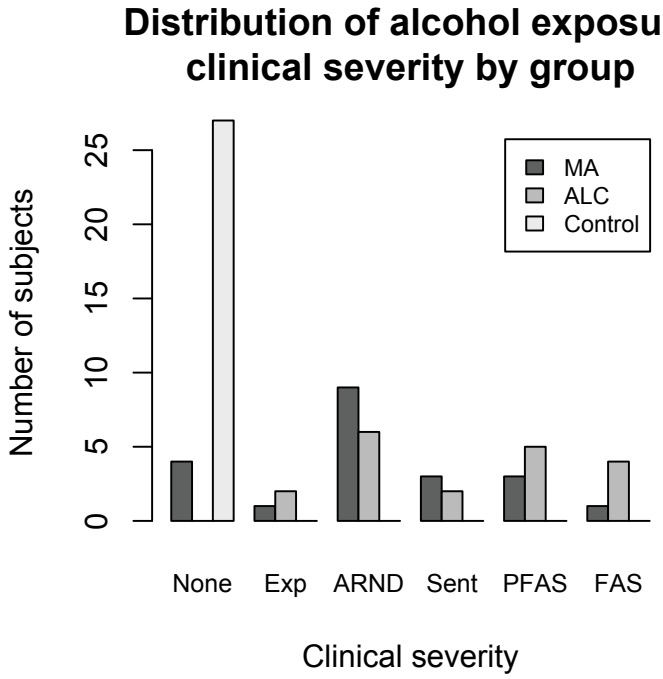


Figure 2.1: Alcohol exposure clinical severity by group. Exp = Exposed (least severe), ARND = Alcohol-related neurodevelopmental disorder, Sent = Sentinel (shows mild facial dysmorphology), PFAS = partial FAS, FAS = Fetal alcohol syndrome (most severe). MA = Methamphetamine-exposed group, ALC = alcohol-exposed group.

of executive function, and Part B, in which the subject must rapidly connect encircled letters and numbers that have been irregularly placed on a sheet of paper, has been shown to be particularly sensitive to cognitive flexibility (Kortte *et al.*, 2002).

2.3.3 DTI acquisition and processing

Diffusion-weighted imaging data were acquired on a 1.5 T Siemens Sonata MRI scanner with six diffusion encoding gradient directions ($b = 1000\,\mathrm{s/mm^2}$), and one non-diffusion-weighted volume ($b = 0$), per acquisition sequence. Two to four whole-brain acquisitions were obtained for each subject (50 axial slices, slice thickness 3 mm, field of view 192 mm, in-plane matrix 64×64, resulting in $3 \times 3 \times 3$ mm isotropic voxels). Brain volumes were skull stripped and a 12 parameter affine registration to the first $b = 0$ volume was applied to correct for eddy current distortions and minor head motion between the acquisition of consecutive diffusion weighted volumes. The

entire DTI sequence was rejected if any of the raw scans contained dropped slices (commonly from ballistic motion during the EPI acquisition of a single volume), but no additional threshold on motion was employed at this stage. A voxelwise diffusion tensor model was fit to the data, and scalar invariant maps were generated of fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), and radial diffusivity (RD). DTI preprocessing was performed using the FMRIB Software Library (FSL) 4.1.0 analysis suite (Smith *et al.*, 2004; Woolrich *et al.*, 2009) (`http://www.fmrib.ox.ac.uk/fsl`), and this workflow was automated with the LONI Pipeline (Rex *et al.*, 2003) (`http://pipeline.loni.ucla.edu`).

Tract-based Spatial Statistics (TBSS) was then used to investigate regional differences in diffusion parameters along the major white matter tracts (Smith *et al.*, 2006b). First, B-spline based nonlinear registration was performed between all subjects' FA maps, and the most representative target subject was chosen by minimizing the overall deformation cost across subjects. This target subject was registered to the ICBM152 1 mm standard template using an affine transformation, and the remaining subjects were brought through this concatenated spatial normalization process. A study-specific mean FA image was generated in standard space, and skeletonised into a tract-based template at an FA threshold of 0.2 (Figure 2.2, green). Each subject's registered FA map was then projected onto this skeleton for voxelwise statistical inference.

2.3.4 Statistical analysis

Group differences in demographics and performance were assessed using analysis of variance (ANOVA), and the Kruskal-Wallis one-way ANOVA for categorical variables. R 2.9.0 (`http://www.r-project.org`) was used for this statistical analysis. Differences in FA were assessed within FSL, and whole-skeleton statistical parametric maps (SPMs; *t*-test, two sample, unpaired) were generated for group differences. Subject age was included as a between-group covariate to model variance in FA due to known effects of developmental maturation. Also, since age effects are commonly gender-specific over this age range — owing to differential timings of puberty and hormonal influences between males and females — potential age-gender interactive effects were considered

in our model. Finally, because prenatal alcohol exposure is known to be associated with white matter abnormalities, an additional whole-brain analysis to explore the most specific effects of methamphetamine included alcohol exposure clinical severity as a parameterized between-group covariate. Threshold-free cluster enhancement (TFCE) was used to incorporate neighborhood information around each voxel and up-weight cluster-like structures in the data (Smith and Nichols, 2009). Nonparametric permutation methods were used to generate empiric *P*-values, and the empirically determined null distribution of the maximum test statistic across space allowed us to correct for multiple comparisons by controlling the family-wise error (FWE) rate (Nichols and Holmes, 2002). The Johns Hopkins white matter atlas was used for stereotactic reporting of anatomical location information (Mori *et al.*, 2008). Areas showing significant group differences in FA were used as region of interest (ROI) masks in post hoc analyses to determine whether the observed changes in diffusion directionality (FA) were associated with any changes in total diffusion (MD) within these regions, or the balance of diffusion in the axial (AD) and radial (RD) directions. Mean FA values were also extracted from these regions for use in brain-behavior analyses. Multiple regression was used to investigate potential direct effects of behavioral score on FA within these regions, independent of both age and the group effect that was modeled in the whole-skeleton analysis. Group-by-score interaction effects were also investigated.

2.4 Results

2.4.1 Demographics

Groups did not differ significantly in age, sex, handedness, parental education, parental IQ, family income, Trails B performance, or number of scan averages. However, the groups did significantly differ in IQ score ($F_{2,58} = 14.85$, $P < 0.0001$), VMI score ($F_{2,62} = 7.08$, $P < 0.005$), birth weight ($F_{2,56} = 10.74$, $P < 0.0005$), adoption rate ($H = 52.1$, $P < 0.0001$), and nicotine exposure rate ($H = 44.3$, $P < 0.0001$). Compared to controls, the MA and ALC groups both had lower IQ and VMI scores, and higher rates of adoption and nicotine exposure. This information is summarized in

Table 2.1. Nicotine exposure histories were unavailable for 8 MA subjects and 10 ALC subjects, and 17 of the MA subjects also had prenatal alcohol exposure. Concurrent psychiatric diagnoses present in the MA and ALC groups included attention deficit hyperactivity disorder (13 MA subjects, 14 ALC subjects), bipolar disorder (4, 3), major depression (2, 1), Asperger syndrome (1, 0), autism (1, 0), pervasive developmental disorder — not otherwise specified (1, 0), schizophrenia (1, 0), and obsessive compulsive disorder (0, 1).

2.4.2 White matter microstructure

Group differences in FA that were identified with the TBSS analysis are summarized in Figure 2.2 and Table 2.2. Additionally, while the statistical testing described below was carried out in a voxelwise fashion, regional summaries collapsed across rough ROIs are also provided in Figure 3 in order to visualize the general relationship between diffusion parameters and to check for gross outliers. Note: Explicitly modeling age-gender interaction effects or IQ did not change the regional patterns of group differences for any of the group contrasts studied.

2.4.2.1 MA vs. CON

FA was significantly higher in the MA group than the control group in the genu of the corpus callosum, left hemisphere internal and external capsules, and corona radiata ($P < 0.05$, FWE-corrected across skeleton) (Figure 2.2A). Within the left anterior corona radiata (LCR), a prominent region showing group effects in the whole-brain FA analysis, we observed a lower overall magnitude of diffusion in the MA group, as measured by MD ($P < 0.05$, FWE-corrected across ROI), as well as lower radial diffusivity (RD; $P < 0.001$). Axial diffusivity (AD) was greater in this area in the MA group than the control group with marginal significance ($P = 0.05$). There were no significant CON>MA differences in FA. This general pattern of group differences remained when alcohol exposure clinical severity (Figure 2.1) was included in the model, localizing particularly robustly to the left corticospinal tract along its entirety ($P < 0.01$) (Figure 2.2C).

Table 2.1: Demographics and performance data by group

	MA (n=21)		ALC (n=19)		CON (n=27)		Group Differences[e]
	Mean	SD[d]	Mean	SD	Mean	SD	
Age (years)	9.76	1.84	10.79	2.32	10.30	3.35	-
Male:Female	13:8	-	11:8	-	11:16	-	-
Birth weight (g)	3290.1	409.4	2592.5	797.7	3355.6	357.5	$P < 0.0005$ (F = 10.74, df = 2,56)
							CON>ALC: $P < 0.05$ ($t = 3.17$, df ≈ 13.1)
							MA>ALC: $P < 0.05$ ($t = 2.82$, df ≈ 14.4)
Handedness (-100 left to 100 right)	67.45	29.60	55.56	62.14	64.12	32.88	-
Prorated full-scale IQ	99.29	12.94	85.11	13.65	109.82	15.88	$P < 0.0001$ (F = 14.85, df = 2,58)
							CON>MA: $P < 0.05$ ($t = 2.39$, df ≈ 40.0)
							CON>ALC: $P < 0.0001$ ($t = 5.29$, df ≈ 37.9)
							MA>ALC: $P < 0.005$ ($t = 3.31$, df ≈ 35.4)
Parent education (years)	15.13	2.29	17.00	1.94	16.33	2.87	-
Parent IQ	108.36	9.09	115.82	7.47	110.69	17.67	-
Family annual income[a]	7.27	2.09	7.19	2.32	6.86	2.75	-
Parent type (Adoptive:Biological)	18:3	-	18:1	-	0:27	-	$P < 0.0001$ (H = 52.1, df = 2)
							MA>CON: $P < 0.0001$ (D = 0.86)
							ALC>CON: $P < 0.0001$ (D = 0.95)
Nicotine exposure? (Yes:No:Unknown)	12:1:8	-	9:0:10	-	0:27:0	-	$P < 0.0001$ (H = 44.3, df = 2)
							MA>CON: $P < 0.0001$ (D = 0.92)
							ALC>CON: $P < 0.0001$ (D = 1)
Trails B (total time)[b]	124.79	65.76	143.39	73.81	140.75	114.06	-
VMI (raw score)[c]	21.29	3.18	21.26	3.97	24.72	3.61	$P < 0.005$ (F = 7.08, df = 2,62)
							CON>MA: $P < 0.005$ ($t = 3.43$, df ≈ 43.9)
							CON>ALC: $P = 0.01$ ($t = 2.97$, df ≈ 36.8)

[a]Ordinal scale, 1 = <$5,000, 2 = $5000-9999, 3 = $10,000-19,999, 4 = $20,000-29,999, 5 = $30,000-39,999, 6 = $40,000-49,999, 7 = $50,000-74,999, 8 = $75,000-100,000, 9 = >$100,000. [b]Trails B = Trail Making Test, part B. [c]VMI = Visuomotor Integration. [d]SD = standard deviation. [e]ANOVA omnibus F-test reported for group differences (Significant pairwise *t*-tests using the Holm modified Bonferroni correction and non-pooled variance are also reported for significant omnibus tests). Kruskal-Wallis ANOVA omnibus H test and Kolmogorov–Smirnov (K-S) pairwise D statistics used for categorical variables. df = degrees of freedom.

2.4.2.2 ALC vs. CON

Although a general pattern of lower FA was observed, no effect survived correction for multiple comparisons.

2.4.2.3 MA vs. ALC

In a direct contrast of the MA subjects with the ALC subjects, the MA group showed significantly higher FA in frontal and temporal areas bilaterally — most prominently in the right external capsule ($P < 0.05$, FWE-corrected across skeleton) (Figure 2.2B). This general pattern also remained during a post-hoc analysis on a small subset of 13 MA and 13 ALC subjects who matched exactly on alcohol exposure clinical severity. There were no significant ALC > MA voxels.

2.4.3 Brain–behavior relationships

Multiple regression analysis, using performance on a measure of frontal executive functioning (Trail Making Test, part B, total time), group membership, group-by-score interaction terms, and age to predict FA in the LCR, failed to reveal any significant interaction between group and Trails B score in predicting FA in this region ($F_{2,52} = 3.15$, $P = 0.051$) or a direct relationship between score and FA ($F_{1,52} = 3.36$, $P = 0.073$). Within the right external capsule (REC) area that distinguished ALC from MA subjects, multiple regression analysis (using group, group-by-score, and age predictors) failed to demonstrate any group-by-score interactive effects ($F_{2,58} = 0.93$, $P = 0.40$), but did demonstrate a significant across-group contribution of VMI performance towards predicting FA ($F_{1,58} = 13.26$, $P < 0.001$). The corresponding across-group partial regression coefficient between FA in the REC and VMI raw score was significantly positive ($b = 2.23 \times 10^{-3}$, $t_{60} = 2.08$, $P < 0.05$). ROI placement, FA distributions, and partial regression plot are included in Figure Figure 2.4.

Table 2.2: Summary of anatomical differences in FA

Contrast	Cluster Index	Cluster size (voxels)[a]	Location[b]	Hemisphere	Coordinates (mm)[c]	*P* value (corrected)
MA>CON	1	4556	Internal capsule (anterior limb)	L	-18, -2, 18	0.02
			- Superior fronto-occipital fasciculus	L	-20, 5, 21	0.02
			- Internal capsule (posterior limb)	L	-8, -8, 0	0.03
			- Cingulum	L	-20, -40, -4	0.03
			- Cerebral peduncle	L	-16, -22, -7	0.03
	2	1395	Anterior corona radiata	L	-16, 34, 10	0.03
			- Corpus callosum (genu)	-	-16, 31, 13	0.03
	3	526	Posterior corona radiata	L	-24, -36, 39	0.04
CON>ALC			None significant			
MA>ALC	1	3744	Inferior longitudinal fasciculus	R	37, -17, -9	0.02
			- External capsule	R	33, 5, 4	0.02
	2	3689	External capsule	L	-26, 4, 14	0.03
			- Inferior longitudinal fasciculus	L	-38, -12, -16	0.03
	3	1531	Cerebral peduncle	L/R	7, -26, -20	0.03
			- Internal capsule (posterior limb)	L	-18, -16, 2	0.03
			- Corticospinal tract	R	7, -21, -28	0.04

[a]Cluster-forming threshold was $P < 0.05$. Only clusters with greater than 100 voxels are listed. Local peaks in different anatomical structures also included (minimum distance between local peaks was set at 5 mm). [b]Taken from JHU white matter atlas. [c]Coordinates in MNI stereotactic space (x,y,z).

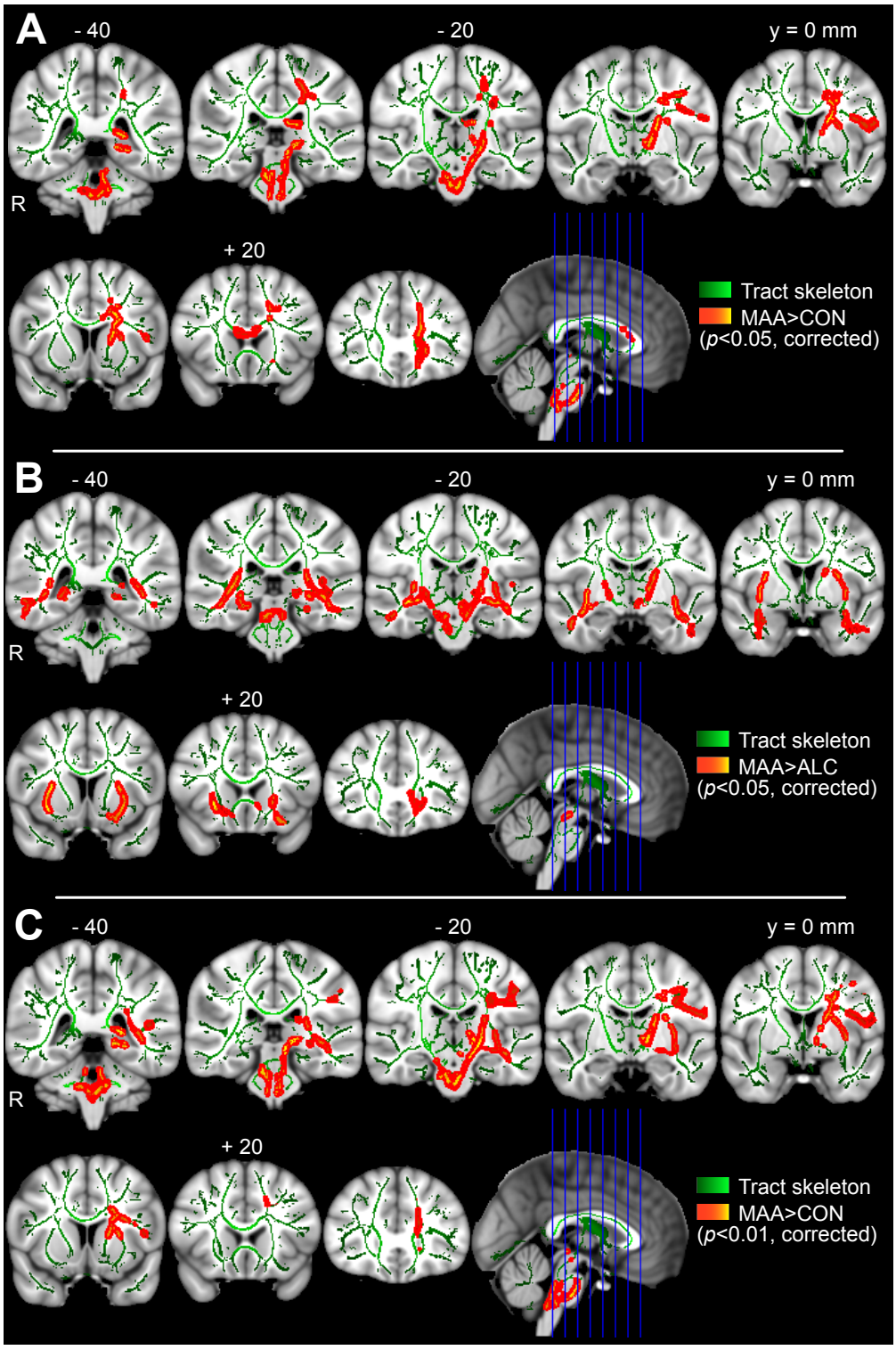

A
- 40
- 20
y = 0 mm
R
+ 20
Tract skeleton
MAA>CON
(p<0.05, corrected)
B
- 40
- 20
y = 0 mm
R
+ 20
Tract skeleton
MAA>ALC
(p<0.05, corrected)
C
- 40
- 20
y = 0 mm
R
+ 20
Tract skeleton
MAA>CON
(p<0.01, corrected)

Figure 2.2: Group differences in fractional anisotropy (FA). A) MA>CON group contrast. B) MA>ALC contrast. C) MA>CON contrast with alcohol exposure clinical severity covariate. Note: Results dilated back into white matter for visualization (red). Areas of greatest significance are displayed as bright centers along the skeleton (yellow and aqua). Background image is ICBM152 1mm standard brain. Slice numbers referenced from y = 0 mm coronal slice in MNI coordinates. MA = methamphetamine-exposed group, ALC = alcohol-exposed group, CON = typically-developing control group.

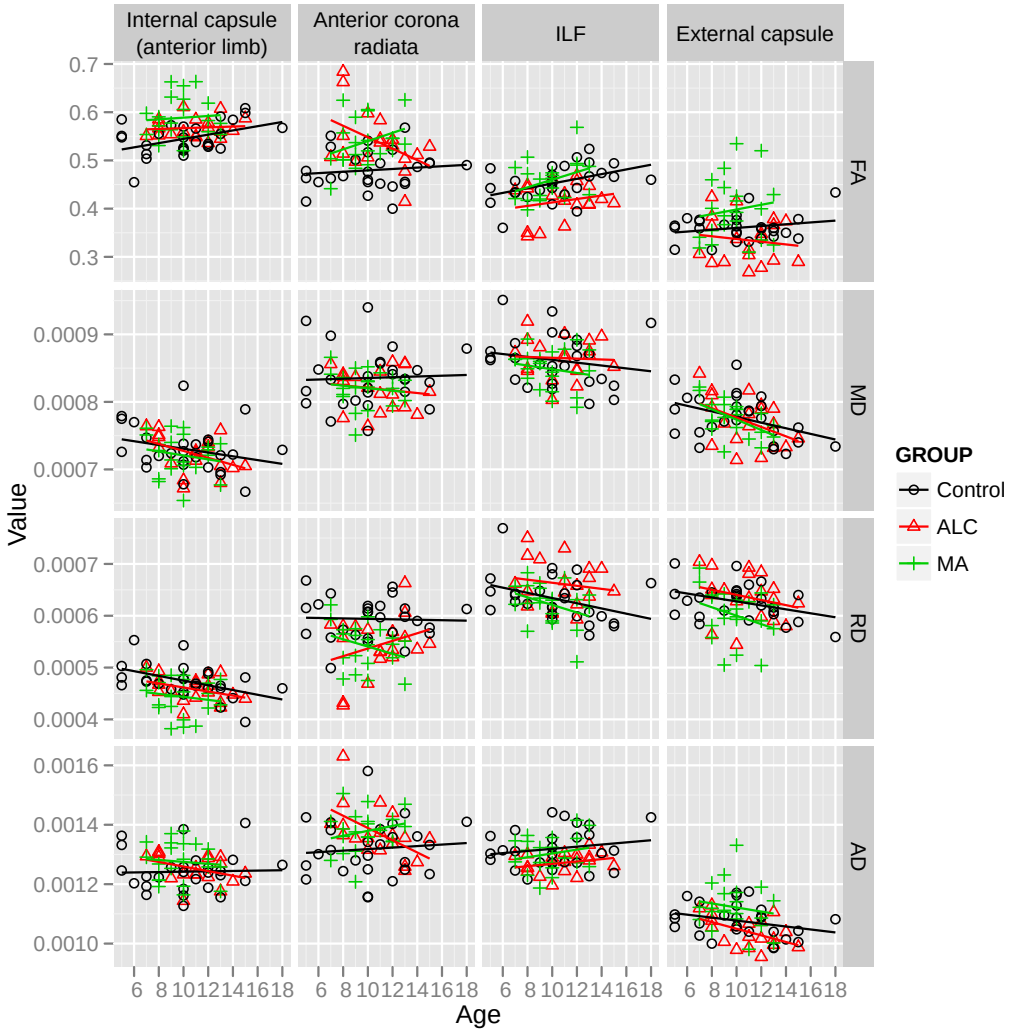


Figure 2.3: Regional scatterplots of DTI metrics. Fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD), are plotted versus age. These plots are facetted into a matrix by brain region, and color is used to encode group membership (Control, ALC, MA). Linear trendlines are added to all plots for visual reference. Each region of interest was obtained by centering a 5mm sphere on the respective cluster center from Table 2, and intersecting this with the associated P<0.05 statistical map from Figure 2. Note: These plots are intended only to give a general regional view of the raw data, and to allow inspection for gross outliers. All statistical testing was performed on the original voxelwise maps.

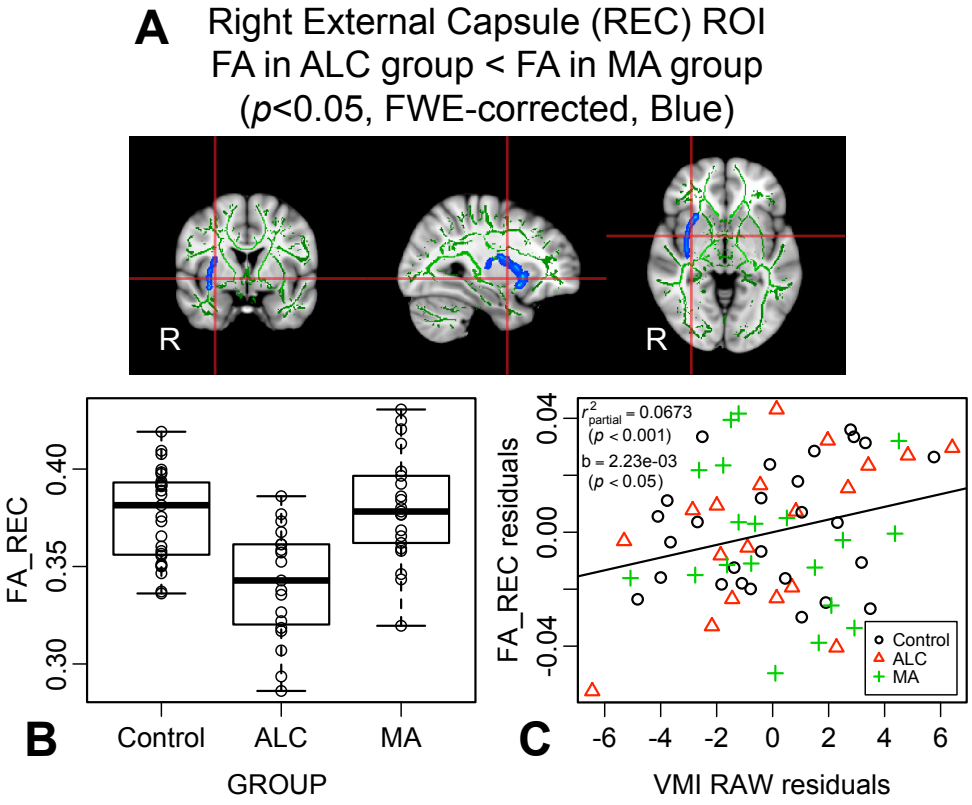


Figure 2.4: Brain-behavior analysis. A) Right external capsule (REC) region of interest (ROI) extracted from thresholded whole-skeleton statistical map (dilated back into white matter for visualization). B) Box and scatter plot displaying the median and quartile distribution of the fractional anisotropy (FA) values for each group within the ROI. C) Partial regression plot between FA in the REC and visuomotor integration (VMI) raw score. Both axes are residualized for age and group.

2.5 Discussion

The MA group, when compared to typically-developing controls, demonstrated higher diffusion anisotropy (FA) in cerebral white matter. Regions showing significant group differences were located mainly along midline structures and in the left hemisphere, and included a pronounced region of the left anterior corona radiata (LCR) (Figure 2.2A). Considering the anatomical connectivity of this tract, this localization is consistent with previous observations of metabolic and volumetric abnormalities in the striata of children with prenatal MA exposure (Chang *et al.*, 2004; Smith *et al.*, 2001), and with the long-standing literature documenting striatal damage in adult MA abusers (McCann *et al.*, 1998). The underlying diffusion pattern accompanying this higher FA — lower RD, higher AD (but to a lesser degree), and, therefore, lower MD — is consistent with a recently published ROI-based DTI study, and an expanded MRS study in the same population, which demonstrated lower diffusion and higher metabolites within small ROIs placed bilaterally in the frontal and parietal lobes of 3 and 4-year-old children with prenatal MA exposure (Chang *et al.*, 2009; Cloak *et al.*, 2009). Further, the authors reported a trend towards higher FA in a left anterior white matter ROI, which also agrees with our results presented here. Our data extend these previous observations into an older age range and with a whole-brain voxelwise analysis, suggesting that these white matter microstructural differences are not short-lived transient effects, but rather broader phenomena that persist later into development. Finally, that the pattern of diffusion differences between MA and CON groups is distinct from the ALC vs. CON contrast, and robustly persists ($P < 0.01$) even when both age-gender interactive effects, and alcohol exposure clinical severity are directly modeled across groups (Figure 2.2C), suggests some level of specificity of prenatal MA exposure towards targeting left-hemisphere white matter regions that connect frontal to striatal structures.

The etiology of changes in FA and other diffusion components in specific brain regions cannot be completely parsed with imaging data (Beaulieu, 2002). Previous reports have loosely associated increased RD with demyelinating disorders (Song *et al.*, 2002), and increased AD with more direct

axonal damage and disruption of the neurofibril architecture (Kim *et al.*, 2006). This framework would suggest that the lower RD we observed may be due to increased myelination of axons within the frontal white matter of children with prenatal MA exposure, a pattern that mimics what is seen during typical development (Lebel *et al.*, 2008b), and would suggest that prenatal MA exposure leads to a pattern of abnormal acceleration in the developmental trajectory of white matter in some brain regions (Bashat *et al.*, 2007; Cloak *et al.*, 2009). If so, it remains to be seen whether this phenomenon represents a direct pathological effect of methamphetamine toxicity, or conversely, a favorable compensatory mechanism in response to insults in other functional systems. Further, because higher FA is not specific to more myelination, higher FA could also represent such scenarios as pathologic decreases in the branching, fanning, or crossing complexities of the neuronal arbor that are manifest through partial volume effects and the uni-orientational nature of the tensor model of diffusion (Silk *et al.*, 2008).

Because our population with prenatal MA exposure also has heavy comorbid exposure to alcohol, we included a separate contrast group of subjects who were exposed to alcohol but not methamphetamine under the rationale that it might serve as a more appropriate real-world control group and allow for better isolation of the specific effects of MA. For instance, nicotine exposure rates and concurrent psychiatric diagnoses — especially ADHD (Fryer *et al.*, 2007) — are better matched between MA and ALC, than between MA and Control groups. In the ALC group, relative to the CON group, we observed sub-threshold trends towards lower FA in the external capsule and deep temporal white matter in the right hemisphere. Still, these results are important to discuss here because they give context to the MA vs. ALC differences: In this direct MA vs. ALC contrast between exposed groups (Figure 2.2B), the MA group demonstrated significantly higher FA than the ALC group in several regions. However, this group effect was most prominent in the right external capsule — a region where the MA group's FA was similar to controls (Figure 2.4B) — suggesting that lower FA in the ALC group is driving this region of highest structural resolvability between the two exposed groups. This is intriguing because many of the MA subjects have also been exposed to alcohol, and yet they do not show this pattern of lower FA. It might suggest that

there are interactive effects with methamphetamine or other factors present in the MA group, or possibly a different pattern of alcohol exposure in the MA group that is beyond what is captured in the relatively matched clinical severity scores.

In order to investigate the possible clinical significance of these structural differences, we related differences in FA to performance on relevant neurobehavioral tests. In line with a previous volumetric study that included cognitive correlates, we failed to demonstrate a significant group difference in the relationship between Trails B performance and white matter in brain regions that differentiated groups (Chang *et al.*, 2004). Further, we failed to observe any relationship between performance on this executive functioning test and FA in the frontal LCR region. When we investigated the relationship between FA in the right external capsule and VMI, we observed a significant relationship between score and FA, independent of baseline group differences in score, but failed to find any group-score interactive effects that would suggest this relationship is modified by exposure status. While this supports a structure-function relationship across groups in this region of cortical connectivity, the specificity of this relationship is questionable — as similar results may have been found in the left external capsule or other tracts relevant to visuomotor performance that were not investigated in this report. However, this result is similar to observations on very low birth weight (VLBW) infants, which include broadly lower FA among VLBW adolescents, and correlations between FA and VMI performance in the internal and external capsules (Skranes *et al.*, 2007). This raises the possibility that there may be interplay between birth weight and alcohol exposure in predicting FA and visuomotor ability. Although it would need to be addressed in a follow-up study containing a low-birth-weight control group, if this were the case it would also help to explain how the MA group has higher FA — since although they share many demographic factors and similar levels of alcohol exposure, the MA group here does not share the lower birth weights of the ALC group.

Several important limitations should be considered when interpreting these results. Because of the clinical population and correlational nature of these findings, influence by confounding variables is always a possibility. To minimize this risk, however, common demographic predictors were

matched across groups, the statistical models were covaried for age, and an alcohol-exposed contrast group was included as a more realistic control group for isolating MA effects. Nevertheless, there could be effects of other factors that correlate specifically with MA use. For example, while we expect nicotine exposure to be relatively well matched between the two exposed groups, direct effects of nicotine could contribute partly to the observed differences between exposed groups and control subjects (Slotkin, 1998). While this study included alcohol exposure clinical severity as a covariate — a novel approach designed to give enhanced specificity for detecting MA effects — likely nonlinearities to this syndrome-FA relationship, as well as potential higher order interaction effects between alcohol and other compounds, remain as possible sources of error. As is true of most retrospective studies of prenatal exposure, precise exposure histories were generally unavailable due to the fact that many of the subjects are not with their biological mothers. Standard limitations of diffusion imaging and the tensor model should also be appreciated. These include artificially depressed FA estimates in regions of complex fiber geometry and partial volume averaging.

As the field transitions from asking the question, "Are there any unique effects of prenatal methamphetamine exposure?" towards actually characterizing these effects, a broadened emphasis is being focused on the integration of observations from different structural, functional, and clinical modalities into a more parsimonious syndromic framework. By identifying a unique pattern of abnormalities in these individuals, we may become better equipped to provide the most appropriate set of behavioral, educational, and occupational interventions to address their specific needs. Importantly, by providing the first independent confirmation of white matter abnormalities in the context of prenatal methamphetamine exposure, by extending the only previous observations into the age range of adolescence and with a whole-brain voxelwise modeling approach, and by evaluating the effects of prenatal MA exposure in the context of an alcohol exposed contrast group, this study helps to solidify the notion that methamphetamine exposure may lead to unique pathological effects on white matter within the developing brain.

CHAPTER 3

National Research Service Award (NRSA) proposal

3.1 Abstract

Fetal alcohol spectrum disorders (FASDs) arise from the teratogenic effects of alcohol exposure, *in utero*, and are an important public health concern as they encompass one of the largest preventable causes of developmental disability. Specifically, deficits in cognitive function are the major contributor to long-lasting morbidity associated with these disorders, and are thought to relate to the broad spectrum of structural brain abnormalities that have also been observed in individuals with FASDs. Executive functioning is one cognitive domain that is particularly affected by fetal alcohol exposure, and includes an important set of cognitive control mechanisms that normally continue to develop through adolescence and help to modulate many other lower order cognitive and motor functions. These actions are thought to be meditated by activity in the frontal lobe, and functional deficits in these areas in children with FASDs may be related to abnormalities in white matter development that have also been observed in the context of fetal alcohol exposure. While frontal lobe white matter maturation also shows a protracted developmental trajectory, and has been related to advancements in executive functioning in typical development, a possible relationship between executive function deficits and white matter abnormalities in FASDs has not been investigated. This proposal aims to apply emerging diffusion magnetic resonance imaging (MRI) techniques, which allow for greater localization and an enhanced focus on developmental timing, to map white matter structural development and its relationship to executive functioning advancements in a sample of typically developing controls. These normative data will then be used as a benchmark for an investigation into the effects of prenatal alcohol exposure on executive function, white matter de-

velopment, and their relationship. Most importantly, the results of these brain mapping efforts will be integrated into a broader classification framework to determine if indices of white matter development can help predict the level of executive function deficit among individuals with FASDs. This work may eventually lead to a better characterization of the precise functional deficits that affect each patient, and facilitate the application of the most appropriate clinical interventions for each individual at an earlier time point than what is allowed by neuropsychological testing alone.

3.2 Specific aims

The study of human brain development is crucial for a better understanding of the emergence of cognitive function in typically developing individuals, and also for the important normative background context that it provides for the study of human disease. In particular, executive functions (*e.g.* working memory, response inhibition, set shifting) continue to develop well into adolescence, and so may be related to the development of white matter connectivity, which also exhibits a uniquely extended developmental trajectory into the more superior and anterior brain regions responsible for higher order cognitive functions.

Fetal alcohol spectrum disorders (FASDs) arise from the teratogenic effects of alcohol exposure, in utero, and are an important public health concern as they encompass one of the largest preventable causes of developmental disability. Children with FASDs exhibit both deficits in executive functioning and white matter abnormalities. Therefore, this population may provide a valuable opportunity to increase our understanding of the interaction between executive function development and white matter development, and to apply this enhanced knowledge towards developing improved clinical interventions for this patient group.

We propose to use emerging diffusion imaging MRI techniques to map white matter structural development in a sample of typically developing controls. We will then build on this foundation by examining the relationships between measures of executive function and white matter development among these individuals. These normative data will then be used as a benchmark for an investi-

gation into the effects of prenatal alcohol exposure on executive function, white matter development, and their relationship. Finally, this information will be integrated into a broader classification framework, incorporating structural imaging and neurobehavioral data, to determine if indices of white matter development can help to accurately classify FASD subjects from controls, and to help predict the level of executive function deficit among the FASD group.

Develop techniques for the study of neuroimaging data in developing populations to…

(a) Evaluate broad trends in the timing of developmental changes in neuroimaging metrics.

(b) Measure localized differences in white matter properties along anatomically defined white matter tracts.

(c) Integrate multimodal imaging (DTI, morphometry, cortical thickness) and behavioral/ demographic data in a classification framework for the prediction of executive function deficits among FASD subjects.

1. **Investigate typical white matter development and its relationship to functional advancements during childhood, adolescence, and young adulthood.**

   (a) **Voxelwise:** We expect to demonstrate a maturing pattern of changes in DTI metrics throughout the brain (increasing fractional anisotropy, FA; decreasing mean diffusivity, MD; decreasing radial diffusivity, RD). We predict broad gradients in the timing of developmental changes in these diffusion metrics, with changes occurring earlier in more posterior/inferior areas, and later in more anterior/superior areas.

   (b) **Tractography:** We expect to demonstrate a maturing pattern of changes in DTI metrics in the major white matter tracts in the brain. We expect that mapping white matter maturation along the tracts will reveal regional variability in the developmental trajectory that characterizes the relationship between diffusion imaging metrics and age (*i.e.* varying age at plateau).

(c) **Brain/behavior:** We expect more efficient processing (*i.e.* higher performance) on measures of executive functioning to be related to more mature DTI indices, independent of age. We expect this relationship to be strongest within tracts that connect to the frontal lobe.

2. **Investigate the impact of prenatal exposure to alcohol on maturational patterns of white matter development, and its relationship with functional deficits.**

   (a) **Voxelwise:** We expect a maturing pattern of changes in DTI metrics throughout the brain in FASD subjects. However, independent of age, we expect altered white matter integrity in FASD subjects. We expect altered global gradients in developmental timing in FASD subjects, relative to patterns in unexposed controls.

   (b) **Tractography:** We expect altered white matter integrity in FASD subjects within major white matter tracts. For areas previously identified to show group differences, we predict that mapping DTI indices along white matter tracts will allow for more precise localization of the affected regions.

   (c) **Brain/behavior:** We expect decreased performance measures in FASD subjects. Within the FASD group, similar to controls, we expect more efficient processing on measures of executive functioning to be related to more mature DTI indices, and expect this relationship to be strongest within tracts in the frontal lobe. While we expect similar patterns of white matter development, we expect the spatial extent of relationships between white matter and executive function to be more diffuse in FASD relative to control subjects, as determined by significant group by score interactions in predicting white matter indices.

   (d) **Classification:** We expect diffusion imaging data to provide useful features for the classification of FASD subjects from controls, and for the prediction of the level of executive function deficits among FASD subjects.

3.3 Investigate typical white matter development and its relationship to functional advancements during childhood, adolescence, and young adulthood. (Specific Aim 1)

3.3.1 Significance

Human brain development is a regionally and temporally dynamic orchestration that begins in utero, and extends well into adulthood. Although overall brain size changes relatively little past the age of 5 (Dekaban, 1978), childhood and adolescence remains a time of remarkable and varied brain maturation beneath the pial surface. Continuing changes include processes that are progressive, like synaptogenesis and myelination, as well as ones that are regressive, like synaptic pruning and programmed cell death (Huttenlocher, 1979). Together, these phenomena help shape the maturation of different cortical and subcortical foci, as well as the white matter network that integrates these diverse regions.

One of the most important, and surprising, observations to emerge from the early developmental neuroimaging efforts to map these maturational changes is that, although gray matter volume (Jernigan *et al.*, 1991; Sowell *et al.*, 2004a), and synaptic density (Huttenlocher, 1979) peak relatively early and then begin to decline during childhood and adolescence, white matter continues a sustained and linear increase in volume through young adulthood (Giedd *et al.*, 1999a; Sowell *et al.*, 2003). Corroborating results come from classical histological studies of myelination, which suggest extended myelination in the white matter well into adulthood (Yakovlev and Lecours, 1967), as well as advancing myelination outwards towards the cortex (Kemper, 1994). This robust and protracted maturational pattern has helped to rewrite the age range associated with brain "development", and has encouraged an increasing research focus on the white matter and its network connectivity as a possible mediator of cognitive gains in executive function domains, which are also seen during later development.

Concurrently with the developing interest in studying white matter connectivity between differ-

ent brain regions and its relationship with cognitive function, diffusion imaging has developed as an MRI modality specifically designed to investigate the white matter (Basser *et al.*, 1994; Beaulieu, 2002; Pierpaoli *et al.*, 1996). Using this unique imaging framework, which provides superior overall white matter contrast compared to traditional T1-weighted anatomical MRI, there has been a recent surge in research aimed at mapping the developmental landscape in these important regions of connectivity. Diffusion tensor imaging (DTI) is the most broadly validated diffusion imaging variation, and fits well within the time constraints of clinical studies (Le Bihan, 2003). This tensor model can be thought of geometrically as a diffusion ellipsoid at each voxel, with diffusion components in the radial (RD, radial diffusivity) and axial (AD, axial diffusivity) directions (Figure 1.6, center). The size of this ellipsoid corresponds to the overall mean diffusivity (MD). The shape of the ellipsoid corresponds to the directionality of diffusion, and is termed fractional anisotropy (FA). By assuming the direction of maximal diffusion aligns with the underlying neuronal fiber geometry, these methods have also been extended with tractography procedures, which allow for the in vivo dissection of groups of tract streamlines that are thought to reflect the underlying white matter tract geometry. Because it has been shown to be sensitive to myelination, FA has received considerable attention as a way to track the developmental maturation within the white matter (Schmithorst and Yuan, 2010), as well as its relationship to cognitive advancements (Bengtsson *et al.*, 2005; Liston *et al.*, 2006; Scholz *et al.*, 2009) and disease (Casey *et al.*, 2007; Le Bihan, 2003). Similar to observations in the overlying gray matter, white matter development — as viewed through the lens of these diffusion imaging metrics — has been shown to be a complex temporally and regionally dynamic function of age (Lebel *et al.*, 2008b; Mukherjee *et al.*, 2001; Snook *et al.*, 2005). The extended period of postnatal myelination contributes to decreasing diffusion in the radial direction (RD) against the myelin insulation, which consequently leads to increasing diffusion directionality (FA). Thus, the characteristic regime of changes seen in DTI metrics during development is that in Figure 1.6D. This period of rapid maturation plateaus towards adulthood, which has led to the preferred use of exponential functions when modeling these changes (Bashat *et al.*, 2005; Lebel *et al.*, 2008b; Mukherjee *et al.*, 2001).

However, despite the strong progress that has been made towards understanding structural white matter development and its functional significance in health and disease, there remain several barriers to progress in this field. Two previous limitations that we help to address with this proposal are: 1) There is commonly a trade-off between the localizability of voxelwise methods, and the personalization (*i.e.* avoidance of spatial-normalization confounds) of tractography methods, and 2) The developmental trajectory is a nonlinear process, characterized by a timing parameter of interest, yet standard voxelwise protocols generally rely on simpler linear modeling methods. To address the first issue, our aims include the use of cutting-edge tractography methods that allow for the analysis of DTI metrics along anatomically defined white matter tracts. To address the second issue, our aims maintain a focus on characterizing the timing of the overall nonlinear developmental trajectory, and how this may relate to functional advancements during typical development, and to disease in the context of fetal alcohol exposure. This is done through the use of a novel "developmental timing quotient", which is sensitive to the nonlinear timing of the developmental trajectory (Aim 1a), and also through the direct modeling of the nonlinear developmental trajectory at many points along tractography-defined tracts (Aim 1b) and the correlation of this timing with executive function advancements (Aim 1c). This emphasis on timing is not typically seen in voxelwise methods, where differential timing is only indirectly inferred from comparisons of binned age groups or regions of interest (ROIs). Yet structural MRI findings in the cortex support the notion that a focus on timing may be of critical importance: In a landmark study, Shaw and colleagues demonstrated that IQ correlates most closely, not with cortical thickness per se, but with the rate of change in cortical thickness during development (Shaw *et al.*, 2006a). It is possible that a similar higher-order timing parameter may be an important way to relate white matter development to functional advancements as well. Therefore, through a combination of improved technical ability and novel analyses, our aims include the first quantitative and continuous analysis of broad trends in developmental timing throughout the brain (Aim 1a), as well as the production of the first along-tract atlases of developmental timing (Aim 1b) and the correlation of this timing with executive function performance (Aim 1c).

Executive functions are a class of cognitive control mechanisms that are thought to be driven by activity in the prefrontal cortex (Fuster, 2000, 2001; Sakai, 2008), and which exert widespread modulatory effects on other lower-order cognitive and motor processes through the postulated sub-domains of shifting, updating (*e.g.* working memory), and inhibition (Miyake *et al.*, 2000; Romine and Reynolds, 2005). These functions are some of the last cognitive domains to fully develop (Luna *et al.*, 2004, 2010; Spear, 2000), which correlates temporally with the protracted course of white matter development. This has prompted the hypothesis that there may be a relationship between these two phenomena, which indeed has been suggested by a recent normal developmental study that correlated response inhibition with average DTI metrics in a tract-based frontostriatal ROI (Casey *et al.*, 2007; Liston *et al.*, 2006), and also by their association in different disease contexts (Skranes *et al.*, 2009). Here we plan to extend these early findings with a more comprehensive mapping of the DTI correlates of executive function (Aim 1c), which will include 1) 10 major white matter tracts in the brain, 2) A broader set of executive functions, and 3) Along-tract procedures for greater localization.

3.3.2 Participants

Participants and recruitment strategies are described more fully, as requested, in the "Protection of Human Subjects" section. Briefly, typically developing participants will include approximately 70 5–30 year old males and females recruited from the Los Angeles community by word of mouth, and by paper flyer, mail, and internet announcements. Exclusion criteria include items that would preclude safe participation in the study, or would be expected to confound results.

3.3.3 Behavioral measures of executive function

Our neuropsychological testing battery includes an estimation of full-scale IQ (WISC-IV), as well as several measures of executive functioning that we intend to correlate with DTI metrics. These include the Wisconsin Card Sorting Test (WCST), a composite test that requires problem solving

and cognitive flexibility, the WISC-IV Working Memory score, and the Delis-Kaplan Executive Function Scale (D-KEFS) battery (Mattson *et al.*, 1999). See Appendix for a full description of these instruments.

3.3.4 MRI data acquisition/processing

MRI data are being collected on a Siemens Trio 3 Tesla scanner at the UCLA Brain Mapping Center as part of Dr. Sowell's ongoing studies of typical development and fetal alcohol exposure. The parts of our acquisition protocol that are relevant to the present proposal are a sagittally-acquired T1-weighted anatomical scan (MP-RAGE), as well as a diffusion weighted series (30 diffusion encoding gradient directions with $b = 1000\,\mathrm{s/mm^2}$, one $b = 0$ volume, 96×96 matrix in-plane, field-of-view 240 mm, 55 axial slices of 2.5 mm each, resulting in $2.5 \times 2.5 \times 2.5$ mm isotropic voxels with whole brain coverage).

The raw diffusion weighted data for each subject will be preprocessed according to standard protocols available in FSL (`http://www.fmrib.ox.ac.uk/fsl`) and TrackVis (`http://www.trackvis.org`), and automated within the LONI Pipeline (`http://pipeline.loni.ucla.edu`). Affine registration to the $b = 0$ volume will be used to correct for minor head motion and eddy-current distortions. A six-parameter tensor model of diffusion will then be fit to the raw data to give voxelwise maps of the 3 principle diffusion directions, as well as the magnitudes of diffusion (*i.e.* eigenvalues) along these three axes. AD is the eigenvalue along the principle eigenvector, and RD is the average of the remaining two eigenvalues. Composite maps will also be generated of FA and MD.

Whole-brain brute-force tractography will be performed using the Fiber Assignment by Continuous Tracking (FACT) algorithm that was developed at Johns Hopkins University (Mori *et al.*, 1999) and has been extensively validated in the literature (Johansen-Berg and Rushworth, 2009; Wakana *et al.*, 2004). This process generates deterministic streamlines by iteratively moving from voxel to voxel along the direction of maximal diffusion. Constraints on the tractography include: 1) A whole-brain mask, 2) An FA threshold of 0.2, to prevent spurious fibers in areas of high directional uncertainty, and 3) A turning angle threshold of 35 or 60 degrees, depending on tract, to

prevent biologically implausible fibers (*e.g.* loops, U-turns, etc.).

3.3.5 Voxelwise modeling of age-related changes in DTI metrics

First-level voxelwise modeling of developmental changes in FA will be conducted for use with the novel developmental timing quotient procedure, discussed next. Methods will follow the standard Tract-based Spatial Statistics (TBSS; part of FSL) pipeline, which attempts to minimize the effects of misregistration errors by collapsing white matter tracts onto a core "skeleton" (Smith *et al.*, 2007). Briefly, this process involves nonlinear registration of FA maps to a standard FA template, and then registration/upsampling into MNI152 standard stereotactic space. The maximal local DTI metrics are then projected down onto the skeleton for voxelwise statistical analysis. Because nonlinear exponential models do not fit within the general linear model (GLM) framework of FSL, for this analysis we model the developmental changes of DTI metrics as a low order polynomial function with age and age^2 regressors. Threshold-free cluster enhancement (TFCE) is used to upweight cluster-like features in the data, and permutation methods are used to generate voxelwise significance *P*-value maps fully corrected for multiple comparisons (Nichols and Holmes, 2002; Smith and Nichols, 2009).

3.3.6 Developmental timing quotient (Aim 1a)

Human brain development is generally thought to proceed along a caudal-rostral arc, with more posterior/inferior primary sensory regions developing earlier than more anterior/superior frontal regions. However, previous observations of patterns in developmental timing of the white matter and myelination have generally been qualitative in nature — either noted through inspection (Kemper, 1994; Yakovlev and Lecours, 1967), or comparisons between binned ROIs (Lebel *et al.*, 2008b; Schneider *et al.*, 2004) or binned age groups (Nomura *et al.*, 1994; Qiu *et al.*, 2008; Snook *et al.*, 2005). Here, we aim to demonstrate this most general phenomenon for the first time in a continuous and quantitative manner across the whole brain.

To do so, we introduce a novel "developmental timing quotient", which uses the *P*-value maps for the age and age^2 regressors from the first-level voxelwise analysis, discussed above. For a given arbitrary ROI, this metric is equal to the fraction of voxels that have a significant linear (*i.e.* age) regression coefficient that additionally have a significant quadratic (*i.e.* age^2) regression coefficient. Intuitively, this can be thought of as similar to "the fraction of the volume that is still showing any age-related increase in FA in our age window, which is nearing its developmental plateau in the neighborhood of our age window." For areas that are developing earlier, a greater proportion of the voxels will exhibit this nonlinear component (*i.e.* will have a significant age^2 term) as they near their developmental plateau, and this timing quotient will be closer to unity. Conversely, areas that are developing later will have a timing quotient closer to zero, as white matter development continues beyond the age range studied and fewer voxels exhibit a significant bending towards a plateau. As such, this metric is expected to be sensitive to the nonlinear developmental timing parameter of interest, yet fits within the simpler GLM framework of standard voxelwise modeling packages.

Using data from an older set of 6-direction DTI scans, we applied this developmental timing quotient in a slice-wise manner to investigate whether it could resolve the expected broad global trends in developmental timing along the inferior-to-superior and posterior-to-anterior axes in the brain. The preliminary results have been promising (See Chapter 4), as this procedure revealed the expected changes in developmental timing, and for the first time, allowed us to quantify these gradients in a continuous manner across the whole brain. This effort has prompted us to extend these efforts to the 30-direction DTI data that we are presently collecting.

3.3.7 Length-parameterized tractography statistics (Aims 1b, 1c)

Traditional voxelwise brain-mapping strategies have revealed that there is prominent regional variation in the characteristics of white matter maturation, as well as their relationship with functional advancements. However, direct voxelwise comparison of diffusion imaging data is challenging because the high-contrast edges of diffusion imaging volumes (*e.g.* FA maps) amplify their susceptibility to small misregistration errors. Further, the anatomical variability of tract position in

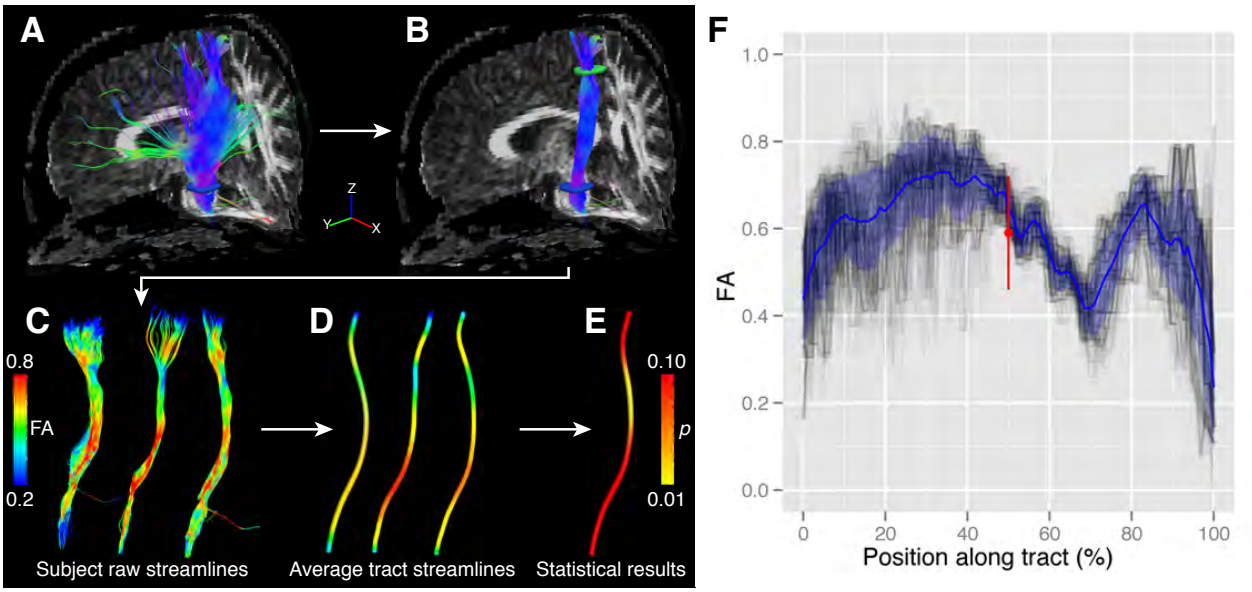


Figure 3.1: Length-parameterized tractography statistics

health, and especially in disease, increases the difficulty of automatically registering functionally equivalent locations (analogous to the difficulty in registering functionally equivalent cortical areas together because of sulcal variability). DTI tractography, on the other hand, is able to circumvent these registration issues by performing individualized tract delineations for each subject, and then makes statistical comparisons by relying on the assumption that the experimenter is actually able to extract the same tract for each subject. Unfortunately, despite the eloquence of this technique, the intricate personalized tract dissections are typically collapsed down in a heavy-handed manner for statistical analyses. This standard processing step ignores the potentially rich anatomical variation in diffusion imaging metrics along the tracts, and yields only a single mean DTI metric and variance estimate for each tract and for each subject (See "collapsed estimate" in Figure 3.1F). Beyond the decreased detail this means for brain mapping studies of typical white matter development, it also implies possibly reduced power to detect group differences within patient populations, as subtle differences may be averaged out across the whole tract, and the variance estimates could be inflated because the potential along-tract variance is ignored.

Accordingly, there is a small but growing interest in modeling DTI parameters along tracts (Corouge *et al.*, 2006; Davis *et al.*, 2009; Zhu *et al.*, 2010). To address this issue, we have developed

a toolbox, written in MATLAB and R, that is able to: 1) Read in dissected tract groups created by the popular TrackVis deterministic tractography software (Figure 3.1A-C), 2) Re-parameterize the included streamlines using cubic B-splines, 3) Resample the streamlines along a uniform set of arc distances from a specified origin, 4) Collapse the raw streamlines from the tract group into an average tract streamline with cross-sectional average DTI metric and variance estimates at many points along the tract (Figure 3.1D,F), 5) perform statistical analysis, including the use of non-linear developmental models, and 6) project the statistical results (parameter estimates and p-values) back onto a representative subject average tract streamline for visualization (Figure 3.1E) (Colby *et al.*, 2012). Preliminary results from an analysis of the corticospinal tract in a small sample of several typically developing individuals demonstrate the feasibility of this approach. Mean DTI metrics were, indeed, found to have prominent variations along the tract, and, as expected, variance estimates were reduced in many areas along the tract (Figure 3.1F).

In the full analysis that will be conducted under this proposal, a full-time trained research assistant will follow the deterministic tractography "cookbook" described in Wakana *et al.* (2007), to extract ten major white matter tracts in the brain for each participant. We will then generate the first along-tract atlases of developmental timing (*i.e.* the time constant characterizing the time-to-plateau).

3.4 Investigate the impact of prenatal exposure to alcohol on maturational patterns of white matter development, and its relationship with functional deficits. (Specific Aim 2)

3.4.1 Significance

Although anecdotal reports in the historical record have long implicated in utero alcohol exposure as a source of congenital disability, the cluster of clinical signs now known as fetal alcohol syndrome (FAS) was not formally characterized until the landmark observations by Jones and colleagues

in Seattle in 1973 (Jones *et al.*, 1973). Since that initial description of the four classic general features of FAS — 1) Confirmed alcohol exposure, 2) Growth deficiency, 3) Facial abnormalities, and 4) Central nervous system manifestations — fetal alcohol exposure has also been associated with other physical abnormalities, as well as a variety of subtler manifestations, which has given rise to the recognition of a more inclusive continuum of fetal alcohol spectrum disorders (FASDs) (American Academy of Pediatrics, 2000). With the expanding recognition of FASDs over the last three decades, the extent of their incidence and morbidity has become startlingly clear: It has been estimated that up to 30% of females in the U.S. who deliver live born infants report alcohol intake during their pregnancy (Ethen *et al.*, 2009), and the incidence of full FAS has been estimated to be between 0.5 and 2 cases per 1000 live births (May and Gossage, 2001). With the inclusion of alcohol-related neurodevelopmental disorder (ARND), an FASD variant that exhibits CNS and cognitive abnormalities but doesn't require the facial stigmata or growth deficiency for diagnosis, this estimate jumps to nearly 1 in every 100 live births (Sampson *et al.*, 1997). This picture becomes even more concerning in select communities where alcohol abuse is endemic (May *et al.*, 2008). Further adding to the expected true impact of FASDs is the fact that underreporting is likely to be significant, due to: 1) Diagnostic difficulty, especially of subtler cases, 2) Likely underreporting of alcohol intake by mothers, and 3) The most severe manifestations of fetal alcohol exposure — spontaneous abortion or still birth — not being reported.

While the relative prominence of physical malformations actually improves with age and maturation through puberty, the neurological manifestations are long lasting into adulthood (Streissguth *et al.*, 1991). Accordingly, a large portion of FASD research has focused on characterizing functional deficits in cognitive, behavioral, and neurological domains, as well as the underlying structural brain abnormalities that are thought to mediate these effects. Neuropsychological testing has revealed a broad range of deficits among FASD subjects, including decreases in full-scale IQ and more specific tests of executive functioning, mathematical ability, motor skills, attention, language, and learning and memory (Gray *et al.*, 2009; Guerri *et al.*, 2009; Howell *et al.*, 2006; Mattson *et al.*, 2010). Executive functioning is affected particularly severely, with reported deficits in response

inhibition, working memory, verbal and nonverbal fluency, planning, attention, cognitive flexibility, feedback utilization, and sequencing (Mattson *et al.*, 1999; Mattson *et al.*, 2010; Rasmussen, 2005). In fact, executive functioning performance is even worse than what would be predicted based on decreased IQ alone (Connor *et al.*, 2000), and is likely one of the domains most affected by fetal alcohol exposure (Mattson *et al.*, 2010). Similarly, in addition to global microcephaly, widespread localized abnormalities in brain structure have also been reported in FASD subjects using neuroimaging methods (Norman *et al.*, 2009; Sowell *et al.*, 2008, 2010; Spadoni *et al.*, 2007). However, of the handful of FASD DTI studies of the white matter, the only clear consensus is on abnormalities in the corpus callosum (Li *et al.*, 2009; Ma *et al.*, 2005; Wozniak *et al.*, 2006). The three remaining studies, which also saw differences in the corpus callosum, but additionally looked elsewhere throughout the brain, are difficult to compare because each used a very different analysis protocol (voxel-based morphometry, TBSS, deterministic tractography) (Fryer *et al.*, 2009; Lebel *et al.*, 2008a; Sowell *et al.*, 2008). Here, we aim to extend previous findings by investigating the effects of fetal alcohol exposure on broad gradients in developmental timing (Aim 2a) and by generating the first along-tract atlas of white matter abnormalities in FASD subjects (Aim 2b). This will advance our knowledge of white matter abnormalities in FASDs by: 1) Extending previous tractography results with far-greater along-tract detail, 2) Giving increased power to detect localized changes within tracts, and 3) Allowing for a link between previous voxelwise and tract-averaged results and a level of corroboration between the studies that have used each method. We will then extend this work with a novel investigation into the relationship between white matter abnormalities and executive functioning deficits (Aim 2c). While both have been reported in the context of FASDs, their relationship has not been investigated. Finally, we will apply the findings from all of these efforts in Specific Aims 1 and 2 to investigate whether DTI data can help to classify FASD subjects from controls, and to predict the level of deficits in executive functioning among FASD subjects (Aim 2d).

3.4.2 Recruitment of FASD subjects

Recruitment strategies and clinical diagnostic procedures for FASD subjects are described more fully, as requested, in the "Protection of Human Subjects" section. In addition to the general recruitment strategies described above for all subjects, participants with prenatal exposure to alcohol were also specifically recruited from a non-clinic-referred university-associated social skills training group for children with FASDs. Subjects with confirmed alcohol exposure were then evaluated for an FASD diagnosis by a skilled clinician, according to the Diagnostic Guide for Fetal Alcohol Syndrome (FAS) and Related Conditions (Astley, 2004). Possible diagnoses include FAS, partial PAS, sentinel features (if not all of the facial features are present for partial FAS), or alcohol-related neurodevelopmental disorder (ARND).

3.4.3 Developmental gradients in FASD subjects (Aim 2a)

We plan to extend our novel characterization of global gradients in white matter developmental timing to subjects with FASD. Based on previous results in the literature, including findings of decreased FA and less efficient functional MRI activation in FASD subjects, we predict altered patterns of developmental timing in FASD subjects. This could manifest as a decreased slope or intercept on the developmental timing quotient vs. slice index plots.

3.4.4 Analysis of DTI metrics along white matter tracts in FASD subjects (Aims 2b, 2c)

The challenges to performing accurate spatial normalization of DTI volumes are compounded in FASD subjects due to gross cranial malformations like microcephaly. This makes in vivo tractography an attractive option for investigating possible abnormalities in white matter maturation in these individuals, as it will be personalized to each subject's individual anatomy. Indeed, the only study that has comprehensively mapped average DTI indices within the major white matter tracts in the brains of FASD subjects was able to demonstrate differences between FASD subjects and controls in a variety of tracts — including the corpus callosum, cingulum, and inferior/superior longitudinal

fasciculi (Lebel *et al.*, 2008a). While these observations helped to extend earlier voxelwise results, it would be useful to localize precisely where along the white matter tracts the differences occur, as well as to investigate differences in developmental timing along these tracts — analyses that have not been previously attempted. By applying the length-parameterized tractography procedure we discussed previously, we aim to create the first along-tract atlas of white matter abnormalities in FASD subjects, and the first mapping of their relationship to executive functioning deficits. This will be done by: 1) Examining group differences at each point along each tract, controlling for age, 2) Examining differences in developmental timing along each tract, and 3) Relating executive function scores to DTI metrics along each tract, controlling for age.

Using this procedure, we have conducted a preliminary analysis in a group of 20 participants (9 with FASDs, and 11 controls). We examined FA along the inferior longitudinal fasciculus (ILF) and the arcuate fasciculus (AF) bilaterally (L and R), and found preliminary group differences in the L ILF that were localized primarily to the posterior section of the tract (See Chapter 5, specifically Figure 6). This finding is in agreement with the results of Lebel et al., but is also a useful illustration of how this procedure provides additional localization detail and the opportunity to reveal features in the data that might be obscured if only whole-tract-averaged DTI metrics are evaluated. These results can then be overlaid on a single subject's average tract streamline for visualization and for comparison with other voxelwise maps (See Chapter 5, specifically Figure 7).

3.4.5 Classification of FASD subjects from controls (Aim 2d)

From a public health standpoint, prevention will always play the primary role in the campaign against the ill effects of fetal alcohol exposure. Nevertheless, postnatal imaging and behavioral studies can also play an important role in the early diagnosis of FASDs, and, most importantly, in the identification of the precise functional deficits in neurological, cognitive, and psychobehavioral domains that affect each individual patient. Specifically, if imaging can provide useful structural indicators of later functional deficits, then this may at some point aid in the implementation of personalized interventions at an earlier age than what is allowed by school-age neurobehavioral

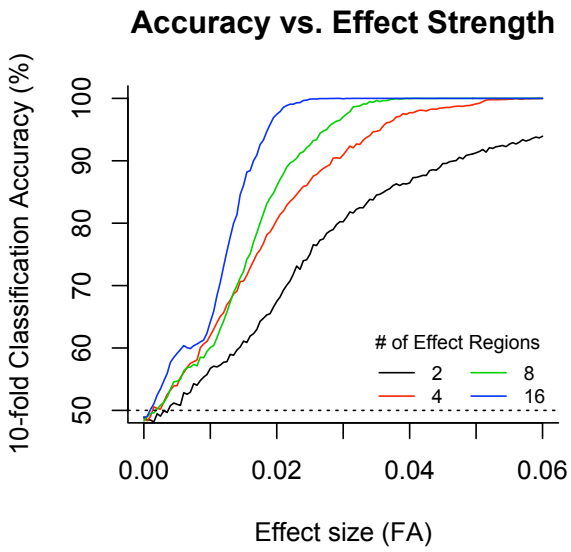


Figure 3.2: Simulated SVM accuracy vs. effect size/extent

testing alone, and may ultimately help mitigate the long-term functional consequences of fetal alcohol exposure. Towards this end, there is growing interest in integrating the findings from various modalities (imaging, clinical findings, neuropsychological testing, demographics) into a common classification system that can be used as a diagnostic aid or for functional prognostic forecasting. Using discriminant function analysis, our group recently demonstrated success (72% overall accuracy, versus 33% chance) using structural morphological features for the classification of FASD and methamphetamine-exposed children from controls (Sowell *et al.*, 2010). A complementary experiment, conducted by our colleagues in the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), of which our lab is also a participating member, focused only on FASD subjects and controls, and used only neurobehavioral data as features. Using an unsupervised latent profile analysis, they were able to confirm that two classes, indeed, fit the data better than one. Using the resultant classifier, they were able to attain an overall accuracy of approximately 85%, versus 50% chance (Mattson *et al.*, 2010). In both studies, the integration of broader features — beyond only full-scale IQ estimates — was able to improve classification accuracy. Here, we aim to extend these efforts by: 1) Integrating structural MRI, diffusion MRI, and neurobehavioral features, and 2) Incorporating support vector machine (SVM) classification and regression, which is a

modern nonlinear method that can also incorporate kernel transformations of the raw data. Using the variance structure extracted from real data, we simulated an artificial dataset to investigate the relationship between effect sizes/extent and classification accuracies. Reports of group differences in FA of 0.02–0.03 across three regions in the context of schizophrenia (Phillips *et al.*, 2009), and 0.03–0.05 across five regions in autism (Kumar *et al.*, 2010) correspond to predicted accuracies of approximately 80–90% (Figure 3.2). This suggests that realistically sized group differences in DTI metrics can result in high levels of group classification accuracy using SVM. We now plan to integrate features from real DTI data with structural imaging and behavioral data to see if this can improve group classification accuracy and/or predict the level of executive function deficits in FASD subjects.

CHAPTER 4

Developmental gradients

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Quantitative *in vivo* evidence for broad regional gradients in the timing of white matter maturation during adolescence

John B. Colby [a,b], John D. Van Horn [a], Elizabeth R. Sowell [a,c,*]

[a] Developmental Cognitive Neuroimaging Group, Laboratory of Neuro Imaging (LONI), Department of Neurology, University of California at Los Angeles, Los Angeles, CA, USA
[b] UCLA Interdepartmental Program for Biomedical Engineering, Los Angeles, CA, USA
[c] UCLA Interdepartmental Program for Neuroscience, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 22 April 2010
Revised 4 August 2010
Accepted 6 August 2010
Available online 12 August 2010

Keywords:
White matter
Myelin
Diffusion
DTI
Development
Gradient

ABSTRACT

A fundamental tenet in the field of developmental neuroscience is that brain maturation generally proceeds from posterior/inferior to anterior/superior. This pattern is thought to underlie the similar timing of cognitive development in related domains, with the dorsal frontal cortices–important for decision making and cognitive control–the last to fully mature. While this caudal to rostral wave of structural development was first qualitatively described for white matter in classical postmortem studies, and has been discussed frequently in the developmental neuroimaging literature and in the popular press, it has *never* been formally demonstrated continuously and quantitatively across the whole brain with magnetic resonance imaging (MRI). Here we use diffusion imaging to map developmental changes in the white matter in 32 typically-developing individuals age 5–28 years. We then employ a novel meta-statistic that is sensitive to the timing of this developmental trajectory, and use this integrated strategy to both confirm these long-postulated broad regional gradients in the timing of white matter maturation *in vivo*, and demonstrate a surprisingly smooth transition in the timing of white matter maturational peaks along a caudal–rostral arc in this cross-sectional sample. These results provide further support for the notion of continued plasticity in these regions well into adulthood, and may provide a new approach for the investigation of neurodevelopmental disorders that could alter the timing of this typical developmental sequence.

Introduction

Much of the current understanding of white matter development has evolved from the seminal work of the pioneering neuroanatomists who first mapped the regional histology of the brain in postmortem samples. Influential observations during the 1960s demonstrated that myelination, while beginning in the second trimester of pregnancy, continues prominently into adulthood through the second and third decades of life—and perhaps beyond (Yakovlev and Lecours, 1967). Moreover, these data extended much earlier observations of regional variations in myelination (Kemper, 1994) and contributed to the wider maxim that brain development generally proceeds along a caudal–rostral arc. Thus, it would appear that myelination within the white matter proceeds in tune with the development of overlying gray matter (Giedd et al., 1999; Gogtay et al., 2004; Huttenlocher, 1990) and the overall emergence of brain function (Spear, 2000)—with more posterior/inferior areas underlying earlier-emerging sensory functions myelinating earlier, and more anterior/superior areas mediating later-emerging higher order executive processes myelinating later.

More recently, MRI has allowed these phenomena to be investigated *in vivo*. In particular, diffusion tensor imaging (DTI) enables measurement of the diffusion properties of water within the brain, which are affected by constraints placed by the neuronal microenvironment (Le Bihan, 2003; Mori and Zhang, 2006). Fractional anisotropy (FA), a measure of diffusion directionality summarizing individual diffusion components along the axial and radial directions (AD and RD, respectively), has been shown to be sensitive to myelination, and so has received significant attention as 1) a metric to track the developmental maturation of white matter (Lebel et al., 2008), 2) a means of mapping normal white matter connectivity (Behrens et al., 2003), 3) a possible imaging biomarker in disease (Sowell et al., 2008; Versace et al., 2008), and 4) a way to investigate the relationship between white matter structure and cognitive function (Bengtsson et al., 2005; Scholz et al., 2009; Tuch et al., 2005). The developmental trajectory of white matter, in general, has been shown to be a complex temporally and regionally dynamic function of age (Hsu et al., 2010; Lebel et al., 2008; Mukherjee et al., 2001). Extended postnatal myelination contributes to a period of increasing FA through greater insulation of the intracellular

* Corresponding author. Department of Neurology, David Geffen School of Medicine at UCLA, 710 Westwood Plaza, Mail Code 176919, Los Angeles, CA 90095-7332, USA. Fax: +1 310 825 6956.
E-mail address: esowell@ucla.edu (E.R. Sowell).

1053-8119/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2010.08.014

environment, a more restricted extracellular milieu, and overall reduced diffusion. As the period of rapid developmental maturation plateaus during adulthood, so too does FA. Accordingly, this nonlinear trajectory has been modeled as an exponential rise (Lebel et al., 2008; Mukherjee et al., 2001), but has also been approximated through young adulthood by polynomial models that integrate well with established whole-brain general linear modeling (GLM) packages (Hsu et al., 2010) (Fig. 1).

Previous developmental DTI reports that have binned different age groups (i.e. child, adolescent, adult) have shown that changes in diffusion parameters are generally concentrated in more posterior occipital regions during the transition between earlier age groups, and in more frontal regions later in development (Asato et al., 2010; Nomura et al., 1994). Similarly, when developmental changes in diffusion parameters have been analyzed with age as a continuous variable, frontal regions-of-interest (ROIs) generally show later timing than other brain regions (Lebel et al., 2008; Mukherjee et al., 2001; Tamnes et al., 2010). Nevertheless, these previous observations of patterns in developmental timing have generally been qualitative in nature—either noted through visual inspection of voxelwise or surface-based statistical maps, or individual pairwise comparisons between binned tract or volume ROIs. Despite the nuanced details about white matter development that have emerged using these strategies, the simplest classical hypothesis–that maturation proceeds in a posterior-to-anterior and inferior-to-superior fashion–remains largely untested in this context.

In the present article, we demonstrate this broad pattern of white matter maturation, as indexed by FA, in the most general sense in a continuous and *quantitative* fashion across the whole brain. To do so, we introduce a novel "developmental timing quotient" that is sensitive to the nonlinear timing of white matter development and fits within a simple GLM framework. While all areas of the brain will eventually reach some sort of developmental peak or plateau, perhaps the majority of regions even within our age window, we reasoned that one could harness the assumed finite power to detect this plateau–usually a limitation of standard analyses–in a useful metric that is sensitive to developmental timing. By doing so, we were able to observe broad gradients in developmental timing within the white matter along posterior-to-anterior and inferior-to-superior axes, and quantify these gradients in a continuous fashion across the brain.

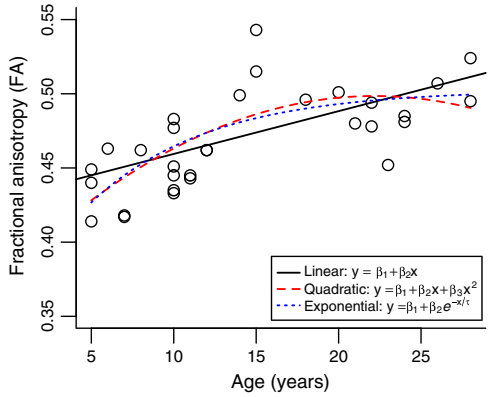


Fig. 1. White matter developmental trajectory example: White matter development, as indexed by FA, deviates from linearity (black line) during adolescence and young adulthood. This has been approximated as a quadratic relationship over this age range (red, dashed), although an exponential fit may be more biologically appropriate because of its stable plateau (blue, dotted). Displayed FA values are taken as the subjects' means from a representative example region in the left corticospinal tract.

Materials and methods

Participants

32 healthy participants (age 5–28 years, mean 14.4 ± 7.2 years (standard deviation, SD), 16 females, 16 males) were recruited in response to advertisements and word-of-mouth. Exclusion criteria included: 1) Known exposure to alcohol or drugs of abuse; 2) age younger than 5 years; 3) IQ less than 70; 4) head injury with loss of consciousness over 20 minutes; 5) physical (e.g. hemiparesis), psychiatric illness, or developmental disability (e.g. autism) that would preclude participation; 6) other potential known causes of mental deficiency (e.g. chromosomal disorders); 7) significant maternal illness with increased risk for fetal hypoxia (e.g. sickle cell disease); and 8) presence of implanted metal in the body. After a thorough discussion of the study protocol, participant consent (or for age<18, parent/guardian consent and participant assent) was obtained in accordance with procedures approved by the UCLA Institutional Review Board.

DTI acquisition and processing

Whole-brain diffusion-weighted imaging data were acquired on a 1.5 Tesla Siemens Sonata MRI scanner. Three sets of whole-brain data were acquired for each subject, with each set including diffusion weighted volumes (6 directions, $b = 1000$ s/mm^2, 50 axial slices, $3 \times 3 \times 3$ mm^3 isotropic voxels) and one non-diffusion-weighted volume ($b = 0$ s/mm^2). Brain volumes were skull stripped with the Brain Extraction Tool (BET) (Smith, 2002) and a 12 parameter affine registration to the first $b = 0$ volume was applied to correct for minor head motion and eddy current distortions introduced by the gradient coils. A diffusion tensor model was fitted to the data in a voxelwise fashion to generate whole-brain maps of fractional anisotropy (FA; the directionality of the diffusion). DTI preprocessing was performed using the FMRIB Software Library (FSL) 4.1.0 analysis suite (http://www.fmrib.ox.ac.uk/fsl) (Smith et al., 2004; Woolrich et al., 2009), and automated within the UCLA LONI Pipeline 4.2.1 processing environment (http://pipeline.loni.ucla.edu) (Rex et al., 2003).

Tract-based Spatial Statistics (TBSS; http://fsl.fmrib.ox.ac.uk/fsl/tbss) was then used to investigate the regional dynamics of age-related maturational changes in this white matter parameter (Smith et al., 2006). First, B-spline based nonlinear registration was performed between all subjects' FA maps and the FMRIB58 58-subject average FA template using the FSL Nonlinear Image Registration Tool (FNIRT). In turn, this FA template was registered to the ICBM152 1 mm standard T1-weighted template using an affine transformation, and all study subjects were brought through this concatenated spatial normalization and up-sampling interpolation process into MNI152 standard space. A study-specific mean FA image was generated in standard space, and skeletonised into a tract-based template at an FA threshold of 0.2. Each subject's registered FA map was then projected onto this skeleton for voxelwise statistical inference.

Statistical analysis

White matter maturation was first modeled as a curvilinear trajectory along this tract-based skeleton template using standard voxelwise GLM tools. Age-related changes in FA were determined in FSL by including age and age^2 as predictors in a whole-brain GLM. One-way *t*-statistic maps (testing for direct and inverse relationships with the age and age^2 terms) where generated for the strengths of the corresponding regression coefficients, and threshold-free cluster enhancement (TFCE) was performed to up-weight cluster-like features in the data by incorporating neighborhood information around each voxel (Smith and Nichols, 2009). Nonparametric permutation testing was used to empirically determine the

null distribution of the maximum test statistic across space (i.e. across the entire TBSS tract skeleton), control the family-wise error (FWE) rate, and generate empiric p-values fully corrected for multiple comparisons (Nichols and Holmes, 2002).

A novel, yet simple, "developmental timing quotient" was then introduced using these whole-skeleton p-value maps of the significance of the regression coefficients: For a given arbitrary region-of-interest (in the present case, a single axial or coronal slice through the tract skeleton) this metric is equal to the fraction of voxels that have a significant ($p<0.05$, corrected) linear (i.e. age) regression coefficient that *additionally* have a significant quadratic (i.e. age^2) regression coefficient. This quotient was calculated for each axial slice along the inferior–superior (IS) axis, and each coronal slice along the posterior–anterior (PA) axis. Using the slice views in Fig. 2a, b as an example (see also Supplemental Movie online), this would be the number of red/yellow voxels divided by the number of underlying blue voxels in each slice. Intuitively, this can be thought of as similar to "the fraction of the volume that is still showing any age-related increase in FA in our age window, which is nearing its developmental plateau in the neighborhood of our age window." For areas that are developing earlier, a greater proportion of the voxels will exhibit this nonlinear component (i.e. will have a significant age^2 term) as they near their developmental plateau, and this timing quotient will be closer to unity. Conversely, areas that are developing later will have a timing quotient closer to zero, as white matter development continues beyond the age range studied and fewer voxels exhibit a significant bending towards a plateau. As such, this metric is sensitive to the nonlinear timing parameter of interest, and yet fits easily within a simpler linear modeling framework. Importantly, this method assumes that, on average, we have less power to detect nonlinear changes (i.e. a significant age^2 term) in regions that plateau later. We therefore expect fewer significant voxels for the quadratic model term that accounts for this effect in the regions of the brain that plateau later, and ultimately a lower developmental timing quotient. The extreme aspects of the tract skeleton were cropped to only include slices with 100 or more skeleton voxels, as the timing quotient becomes unstable near the extreme edges of the brain where the denominator approaches zero. Finally, weighted least squares linear regression was used to investigate how the developmental timing quotient is related to position in the brain along the IS and PA axes, with the number of voxels in each slice (which is inversely related to the expected variance in the timing quotient) used as the weighting function.

Results

As predicted, voxelwise modeling along the tract-based skeleton demonstrated widespread age-related increases in FA in our 32 subject cross-sectional sample ($p<0.05$, corrected). Further, many of these regions where the age regressor was significant also demonstrated a significant negative age^2 term ($p<0.05$), indicating a quadratic-type nonlinear relationship in these regions over our age range. No voxels were observed with a significant inverse relationship between age and FA or a significant positive age^2 term. See slice views in Fig. 2a, b and also the Supplemental Movie online for a full collection of these raw first-level statistical maps.

When the individual diffusion components along the axial and radial directions were analyzed (AD and RD), which together contribute to the FA measure, widespread significant decreases in RD were observed with similar regional and inverse temporal dynamics compared to FA (i.e. a significant negative age term and a significant positive age^2 term were observed; $p<0.05$, corrected) (see Fig. 3a, b). No voxels were observed with significant increases in RD or changes in AD.

By implementing the developmental timing quotient meta-statistic in a slice-wise fashion, we were able to collapse whole-brain three-dimensional statistical maps down into one-dimensional high-level summary views along the inferior-to-superior (Fig. 4a) and posterior-to-anterior axes (Fig. 4b). As long-expected, prominent gradients in the developmental timing of white matter maturation, as measured by FA, were observed and quantified in both the inferior-to-superior ($R^2=0.74$, $b=-0.00614$, $t=-18.36$, $p=1.61e{-}36$, $n=121$ axial slices) and posterior-to-anterior directions ($R^2=0.54$, $b=-0.00456$, $t=-13.35$, $p=2.69e{-}27$, $n=152$ coronal slices). Our results show a dramatic reduction in the volume of voxels that reach their developmental peak in the most anterior and superior regions of the brain, compared to the more posterior and inferior regions, suggesting continued white matter maturation beyond the age range studied here in dorsal frontal brain regions. This general pattern remained unchanged during a *post hoc* analysis that excluded the cerebellum.

Discussion

The observed trends in the developmental timing quotient across the brain provide strong support to the long-held hypothesis that white matter maturation proceeds in an inferior-to-superior and posterior-to-anterior manner, and provide a straightforward confirmation of this fundamental phenomenon in developmental neuroscience. In particular, the results agree with the classical postmortem literature (Kemper, 1994; Yakovlev and Lecours, 1967), while also providing a validation of other neuroimaging studies examining more nuanced details of brain development (Tzarouchi et al., 2009) or aging (Davis et al., 2009) in the context of this often-cited global developmental pattern—one that has not been previously demonstrated in a continuous and quantitative fashion. Importantly, this picture suggesting prolonged myelination in frontal regions also is consistent with the protracted trajectory of cognitive development in executive functioning domains, which similarly continues through adolescence and is known to involve processing in the frontal lobe (Luna et al., 2004, 2010; Romine and Reynolds, 2005). Previous observations have linked both DTI-measured connectivity (Liston

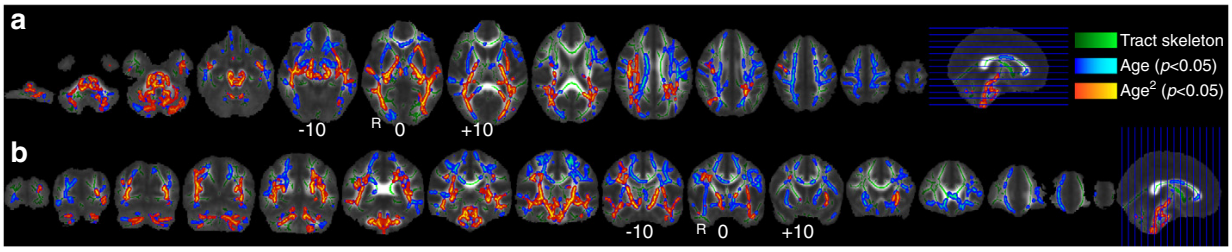


Fig. 2. TBSS voxelwise results: Statistical maps of the significance of the age (blue-aqua, $p<0.05$, corrected) and age^2 (red-yellow, $p<0.05$, corrected) regressors, from the first-level tract-based spatial statistics (TBSS) analysis of age-related changes in fractional anisotropy (FA), are overlaid on the white matter tract skeleton (green) and the study-specific mean FA template volume. Axial (a) and coronal (b) slices are displayed in 10 mm increments. The complete collection of slices is available as a Supplemental Movie online.

Fig. 3. FA, AD, and RD changes by region: (a) Developmental trajectory curves for FA are shown from a representative "inferior" ROI in the corticospinal tract that was classified as quadratic (colored red to match the red/yellow quadratic voxels in Fig. 2), and a "superior" ROI in the cingulum that was still increasing linearly (colored blue to match the blue linear voxels in Fig. 2). (b) The axial diffusivity (AD) and radial diffusivity (RD) components are displayed for the same regions, demonstrating the general pattern that the observed FA changes are predominantly due to changes in RD. Trendlines are selectively added according to which terms of the model (age or age + age^2) are significant ($p<0.05$, corrected) in the FA, RD, and AD voxelwise analyses.

et al., 2006; Madsen et al., 2010) and fMRI-measured activation patterns (Bunge et al., 2002; Casey et al., 1995, 2005; Luna et al., 2001) with performance on cognitive control tasks across this age range, strongly implicating these systems as part of the neural basis for real-world changes seen during adolescence in areas like risk/reward processing (Chau et al., 2004; Olson et al., 2009; Spear, 2000), and also as a potential contributing factor to neurodevelopmental disorders like ADHD (Casey et al., 2007) and neuropsychiatric illnesses that often begin to emerge during adolescence (Reichenberg et al., 2010; Versace et al., 2008). The implications of this new approach are particularly valuable as they relate to the period of adolescent brain development. Because this technique extracts regional developmental timing information in a *continuous* manner and *across* an age window, it may prove useful to investigators as an attractive means of probing developmental delays and other clinical phenomena that appear during adolescence. For instance, the slope of the developmental timing gradient across an entire region and the entire age window could be compared between groups, as a complementary method to traditional examination of a series of regions-of-interest between several binned age groups in a pairwise manner. In particular, processing speed is an attractive target for further study, as it shares an intuitive foundation in connectivity, and has been shown to exhibit some of the strongest advances during adolescence (Anderson et al., 2001). This may be related, in part, to concomitant prolonged myelination in the white matter through an advancing structural network and increasing transmission efficiency. Further, recent

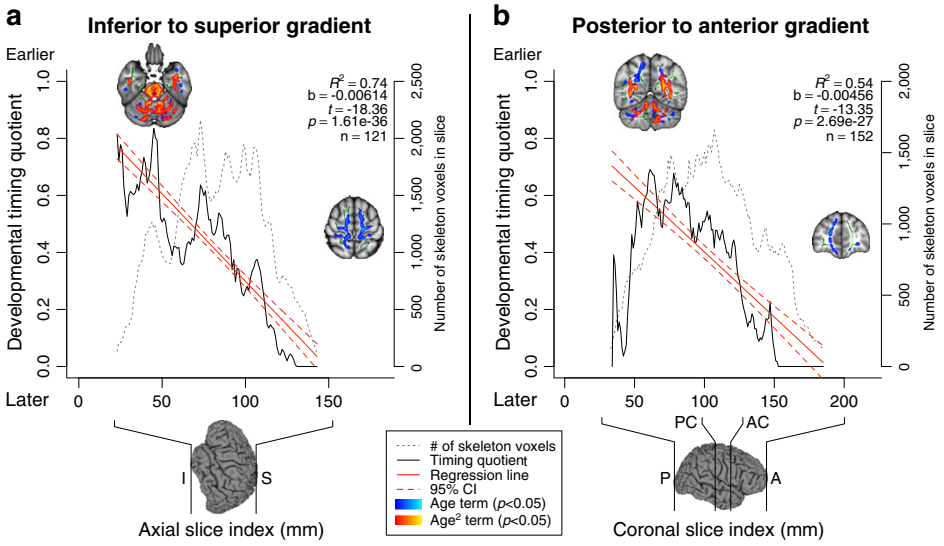


Fig. 4. Gradients in white matter developmental timing: White matter development takes place along inferior-to-superior (a) and posterior-to-anterior (b) gradients. The developmental timing quotient is plotted against slice index (mm), and the weighted least squares linear regression line (± 95% pointwise confidence interval) is overlaid. The number of white matter tract voxels in each slice was used as the weighting function, and is plotted on the right *y*-axis. Slice overlays show representative examples of the statistical maps used to calculate the timing quotient (blue-aqua, age term ($p<0.05$); red-yellow, age^2 term ($p<0.05$); green, white matter tract skeleton). I, inferior; S, superior; P, posterior; A, anterior; PC, posterior commissure; AC, anterior commissure.

evidence demonstrating *slower* advances in processing speed in children preceding adult schizophrenia makes the quantified regional timing of white matter development discussed in this report a particularly attractive metric that may also be slowed in these individuals (Reichenberg et al., 2010). These core developmental gradients could also be investigated more closely as quantifiable imaging biomarkers in the context of typical development. It is possible that, like similar unexpected findings in gray matter (Shaw et al., 2006), a higher-order timing parameter like the one described may explain individual variance in performance better than the raw diffusion imaging parameters at any one point in time.

Although not the primary focus of the present investigation, first-level voxelwise results are in general agreement with the published literature documenting broadly distributed increases in FA during development (Figs. 2, 3a) (Lebel et al., 2008; Mukherjee et al., 2001; Qiu et al., 2008; Snook et al., 2005). Analysis of the underlying axial and radial diffusion components suggests that decreasing RD is the dominant contributor to the observed changes in FA (Fig. 3b), although subthreshold contributions by changes in axial diffusivity are also likely. A similar pattern of decreasing diffusivity, driven predominantly by decreasing RD, is reported in the majority of developmental DTI studies (Schmithorst and Yuan, 2010).

One of the most interesting aspects of the quantified gradients demonstrated here is the smoothness with which the developmental timing of white matter maturation changes across the brain—particularly along the inferior-to-superior axis (Fig. 4a). The similarity of the developmental timing gradients between the inferior-to-superior and posterior-to-anterior axes is also compelling. While not totally unexpected, since the trajectories along the PA and IS axes will naturally be somewhat correlated because more inferior regions of the brain also contain more posterior voxels, this coordination is also consistent with the notion of a continuous caudal–rostral wave in the timing of white matter maturation. For both trajectories, the timing quotient begins near 0.8 and then decreases robustly, but smoothly and with similar slope, to below 0.4. In fact, the only regional locus that may not warrant inclusion in this continuum model is the occipital lobe (Fig. 4b), which exhibits an unexpected drop in the timing quotient. This could be due, in part, to the unstable nature of the timing quotient as one moves to the extreme edges of the brain and the number of skeleton voxels approaches zero. To accommodate this behavior in the statistical analysis, we chose to apply a weighted least-squares fit to the data, which down-weights these unstable slices with very few voxels. However, upon closer inspection, some of these slices through the occipital lobe actually have enough voxels that we would expect the timing quotient estimates to be reliable (i.e. they are above the *ad hoc* 100 voxel threshold, below which the timing quotient can oscillate unstably). Accordingly, this area in particular warrants future examination to determine to what extent this effect is real.

While this is the first time that these broad gradients in developmental timing have been quantified continuously across the white matter, these results are in agreement with the strong tradition of qualitative DTI observations that have previously been made on this topic—particularly concerning the more focused notion that frontal lobe white matter connectivity tends to have the longest course of development, relative to other regions (See Schmithorst and Yuan, 2010 for an excellent review of this developmental DTI literature). When diffusion parameters have been compared between two (or several) binned age groups, the magnitude of the change in these metrics between age groups tends to be greater in frontal regions, especially when using contrasts that span older age ranges—suggesting that frontal regions are undergoing a larger portion of their maturation at later ages, as compared to more posterior regions (Nomura et al., 1994; Qiu et al., 2008; Snook et al., 2005). Similarly, when the timing parameters of exponential models have been compared across ages within binned regional ROIs, frontal regions tend to have a slower rise (i.e. a later time-to-plateau) than more posterior regions (Lebel et al., 2008; Schneider et al., 2004). Our results reported here provide further evidence in support of the presence of a particularly extended developmental trajectory in the frontal lobe, and additionally help extend this notion to suggest a relatively smooth gradient transition in timing from more posterior/inferior regions.

Another key aspect of the present report is the introduction of the simple developmental timing quotient meta-statistic that was used to demonstrate these findings. This metric can be readily computed from the statistical maps generated by standard GLM analyses, and gives insight into the timing of the developmental trajectory in a given ROI (in this case, a specific slice). Previously, in order for timing phenomena to be investigated, a time constant parameter would need to be modeled directly, which requires a nonlinear model–like the exponential function described in Fig. 1–to be fitted in a voxelwise fashion. This comes at a considerable computational cost, and additionally introduces the practical challenge of being incompatible with the major voxelwise modeling packages in the neuroimaging community. These difficulties perhaps explain why the important fundamental notion of a broad caudal–rostral wave of brain maturation, which has been qualitatively and anecdotally referenced quite often, has until now never been directly demonstrated and quantified with MRI. Importantly, however, one must remain aware that the design presented here offers an *indirect* window into the timing properties of this system. As one consequence, the ability to demonstrate these timing gradients could be affected, for example, if the power to detect individual voxelwise relationships in the first-level GLM analysis was increased to the point that a significant quadratic age^2 term could be demonstrated *everywhere* in the brain. In this case, the timing quotient would saturate and the gradients would be obscured, or perhaps only observable with different first-level significant thresholds. Likewise, if the age window were altered to be older, then the number of regions showing any developmental increases in FA would start to decrease as they complete their plateau. Therefore it is important that the study sample and significance thresholds are tuned to the question of interest.

Although it might be assumed that close to all of the voxels in the brain are exhibiting some form of real neurobiological development over the age range from childhood to adulthood, we were only able to demonstrate significant FA-indexed "development" in a subset of these voxels (those with a significant age term, displayed as the extent of underlying blue voxels in Fig. 2). There are several explanations for this, any or all of which could be contributing, including: (1) The voxels have already "matured" in terms of their developmental changes in FA, (2) The voxels are maturing slowly enough that they not only fail to show a significant quadratic age^2 term, but additionally fail to show a significant linear age term, (3) There is finite power to detect developmental changes (e.g. due to acquisition SNR, intersubject variation, misregistration, etc.), even for structures that *are* developing over our age range, (4) Methodological issues are involved—for instance, those associated with the tensor model of diffusion like partial volume averaging within voxels that contain multiple fiber populations, and (5) Although we failed to demonstrate this anywhere in the brain, some voxels may actually exhibit *decreasing* FA during development (Schmithorst et al., 2008). Other methodological limitations should also be considered when interpreting these results. Although data interpolation is a standard component of the TBSS processing pipeline, it can affect the apparent resolution when visually interpreting the resulting statistical maps. In the present case, this means that the effective resolution across the developmental gradients is likely to be much coarser than the 1 mm thick slices on which the data were processed. Additionally, while the observed trends in developmental timing quotient are expected to be relatively robust to intersubject variation in raw diffusion MRI parameters–as the developmental timing quotient effectively

averages out some of this variation–it should be noted that this sample is relatively sparse during the age range of late adolescence. Increasing the number of subjects studied, especially during this period, should increase power to detect both voxelwise and timing quotient changes, and improve confidence in the shape of the underlying developmental trajectory.

Conclusions

In summary, we have employed a simple quantitative approach using diffusion imaging data in a developmental sample which (1) provides confirmation, in a continuous and quantitative manner across the whole brain, that broad regional gradients are present in the developmental timing of white matter maturation, (2) introduces a novel meta-statistic that is sensitive to the developmental timing of white matter maturation in our study population and fits within the established GLM framework of whole-brain mapping strategies, and (3) presents a new opportunity for the investigation of the structural basis of advances in executive function during adolescence and their relationship to the emergence of neuropsychiatric disease.

Supplementary data to this article can be found online at doi: 10.1016/j.neuroimage.2010.08.014.

Acknowledgments

This research was supported by the following organizations: National Institute of Child Health and Human Development (R01 HD053893), National Center for Research Resources (U54 RR021813), National Institute of General Medical Sciences (T32 GM008042), National Institute on Drug Abuse (R01 DA017831, R90 DA023422), and March of Dimes (6-FY2008-50). The authors report no conflicts of interest.

References

Anderson, V.A., Anderson, P., Northam, E., Jacobs, R., Catroppa, C., 2001. Development of executive functions through late childhood and adolescence in an Australian sample. Dev. Neuropsychol. 20, 385–406.

Asato, M.R., Terwilliger, R., Woo, J., Luna, B., 2010. White matter development in adolescence: A DTI study. Cereb. Cortex 20, 2122–2131.

Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C. A.M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A. J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. 6, 750–757.

Bengtsson, S.L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F., 2005. Extensive piano practicing has regionally specific effects on white matter development. Nat. Neurosci. 8, 1148–1150.

Bunge, S.A., Dudukovic, N.M., Thomason, M.E., Vaidya, C.J., Gabrieli, J.D.E., 2002. Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. Neuron 33, 301–311.

Casey, B.J., Cohen, J.D., Jezzard, P., Turner, R., Noll, D.C., Trainor, R.J., Giedd, J., Kaysen, D., Hertz-Pannier, L., Rapoport, J.L., 1995. Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. Neuroimage 2, 221–229.

Casey, B.J., Tottenham, N., Liston, C., Durston, S., 2005. Imaging the developing brain: what have we learned about cognitive development? Trends Cogn. Sci. 9, 104–110.

Casey, B.J., Epstein, J.N., Buhle, J., Liston, C., Davidson, M.C., Tonev, S.T., Spicer, J., Niogi, S., Millner, A.J., Reiss, A., Garrett, A., Hinshaw, S.P., Greenhill, L.L., Shafritz, K.M., Vitolo, A., Kotler, L.A., Jarrett, M.A., Glover, G., 2007. Frontostriatal connectivity and its role in cognitive control in parent–child dyads with ADHD. Am. J. Psychiatry 164, 1729–1736.

Chau, D.T., Roth, R.M., Green, A.I., 2004. The neural circuitry of reward and its relevance to psychiatric disorders. Curr. Psychiatry Rep. 6, 391–399.

Davis, S.W., Dennis, N.A., Buchler, N.G., White, L.E., Madden, D.J., Cabeza, R., 2009. Assessing the effects of age on long white matter tracts using diffusion tensor tractography. Neuroimage 46, 530–541.

Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., Zijdenbos, A., Paus, T., Evans, A.C., Rapoport, J.L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. Nat. Neurosci. 2, 861–863.

Gogtay, N., Giedd, J.N., Lusk, L., Hayashi, K.M., Greenstein, D., Vaituzis, A.C., Nugent, T.F., Herman, D.H., Clasen, L.S., Toga, A.W., Rapoport, J.L., Thompson, P.M., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. Proc. Natl. Acad. Sci. U.S.A. 101, 8174–8179.

Hsu, J., Hecke, W.V., Bai, C., Lee, C., Tsai, Y., Chiu, H., Jaw, F., Hsu, C., Leu, J., Chen, W., Leemans, A., 2010. Microstructural white matter changes in normal aging: a diffusion tensor imaging study with higher-order polynomial regression models. Neuroimage 49, 32–43.

Huttenlocher, P.R., 1990. Morphometric study of human cerebral cortex development. Neuropsychologia 28, 517–527.

Kemper, T.L., 1994. Neuroanatomical and neuropathological changes during aging, In: Albert, M.L., Knoefel, J.E. (Eds.), Clinical Neurology of Aging, 2nd ed. Oxford University Press, New York, NY, pp. 8–9.

Le Bihan, D., 2003. Looking into the functional architecture of the brain with diffusion MRI. Nat. Rev. Neurosci. 4, 469–480.

Lebel, C., Walker, L., Leemans, A., Phillips, L., Beaulieu, C., 2008. Microstructural maturation of the human brain from childhood to adulthood. Neuroimage 40, 1044–1055.

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. Cereb. Cortex 16, 553–560.

Luna, B., Thulborn, K.R., Munoz, D.P., Merriam, E.P., Garver, K.E., Minshew, N.J., Keshavan, M.S., Genovese, C.R., Eddy, W.F., Sweeney, J.A., 2001. Maturation of widely distributed brain function subserves cognitive development. Neuroimage 13, 786–793.

Luna, B., Garver, K.E., Urban, T.A., Lazar, N.A., Sweeney, J.A., 2004. Maturation of cognitive processes from late childhood to adulthood. Child Dev. 75, 1357–1372.

Luna, B., Padmanabhan, A., O'Hearn, K., 2010. What has fMRI told us about the development of cognitive control through adolescence? Brain Cogn. 72, 101–113.

Madsen, K.S., Baaré, W.F.C., Vestergaard, M., Skimminge, A., Ejersbo, L.R, Ramsøy, T.Z., Gerlach, C., Akeson, P., Paulson, O.B., Jernigan, T.L., 2010. Response inhibition is associated with white matter microstructure in children. Neuropsychologia 48, 854–862.

Mori, S., Zhang, J., 2006. Principles of diffusion tensor imaging and its applications to basic neuroscience research. Neuron 51, 527–539.

Mukherjee, P., Miller, J.H., Shimony, J.S., Conturo, T.E., Lee, B.C., Almli, C.R., McKinstry, R. C., 2001. Normal brain maturation during childhood: developmental trends characterized with diffusion-tensor MR imaging. Radiology 221, 349–358.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.

Nomura, Y., Sakuma, H., Takeda, K., Tagami, T., Okuda, Y., Nakagawa, T., 1994. Diffusional anisotropy of the human brain assessed with diffusion-weighted MR: relation with normal brain development and aging. AJNR Am. J. Neuroradiol. 15, 231–238.

Olson, E.A., Collins, P.F., Hooper, C.J., Muetzel, R., Lim, K.O., Luciana, M., 2009. White matter integrity predicts delay discounting behavior in 9- to 23-year-olds: a diffusion tensor imaging study. J. Cogn. Neurosci. 21, 1406–1421.

Qiu, D., Tan, L., Zhou, K., Khong, P., 2008. Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: Voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. Neuroimage 41, 223–232.

Reichenberg, A., Caspi, A., Harrington, H., Houts, R., Keefe, R.S.E., Murray, R.M., Poulton, R., Moffitt, T.E., 2010. Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: A 30-Year study. Am. J. Psychiatry 167, 160–169.

Rex, D.E., Ma, J.Q., Toga, A.W., 2003. The LONI Pipeline processing environment. Neuroimage 19, 1033–1048.

Romine, C.B., Reynolds, C.R., 2005. A model of the development of frontal lobe functioning: findings from a meta-analysis. Appl. Neuropsychol. 12, 190–201.

Schmithorst, V., Yuan, W., 2010. White matter development during adolescence as shown by diffusion MRI. Brain Cogn. 72, 16–25.

Schmithorst, V.J., Holland, S.K., Dardzinski, B.J., 2008. Developmental differences in white matter architecture between boys and girls. Hum. Brain Mapp. 29, 696–710.

Schneider, J.F.L., Il'yasov, K.A., Hennig, J., Martin, E., 2004. Fast quantitative diffusion-tensor imaging of cerebral white matter from the neonatal period to adolescence. Neuroradiology 46, 258–266.

Scholz, J., Klein, M.C., Behrens, T.E.J., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. Nat. Neurosci. 12, 1367–1368.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., Giedd, J., 2006. Intellectual ability and cortical development in children and adolescents. Nature 440, 676–679.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17, 143–155.

Smith, S., Nichols, T.E., 2009. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. Neuroimage 44, 83–98.

Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., Luca, M.D., Drobnjak, I., Flitney, D.E., Niazy, R.K., Saunders, J., Vickers, J., Zhang, Y., Stefano, N.D., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23, S208–S219.

Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.

Snook, L., Paulson, L., Roy, D., Phillips, L., Beaulieu, C., 2005. Diffusion tensor imaging of neurodevelopment in children and young adults. Neuroimage 26, 1164–1173.

Sowell, E.R., Johnson, A., Kan, E., Lu, L.H., Van Horn, J.D., Toga, A.W., O'Connor, M.J., Bookheimer, S.Y., 2008. Mapping white matter integrity and neurobehavioral correlates in children with fetal alcohol spectrum disorders. J. Neurosci. 28, 1313–1319.

Spear, L.P., 2000. The adolescent brain and age-related behavioral manifestations. Neurosci. Biobehav. Rev. 24, 417–463.

Tamnes, C., Østby, Y., Fjell, A., Westlye, L., Due-Tønnessen, P., Walhovd, K., 2010. Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. Cereb. Cortex 20, 534–548.

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. Proc. Natl. Acad. Sci. U.S.A. 102, 12212–12217.

Tzarouchi, L.C., Astrakas, L.G., Xydis, V., Zikou, A., Kosta, P., Drougia, A., Andronikou, S., Argyropoulou, M.I., 2009. Age-related grey matter changes in preterm infants: an MRI study. Neuroimage 47, 1148–1153.

Versace, A., Almeida, J.R.C., Hassel, S., Walsh, N.D., Novelli, M., Klein, C.R., Kupfer, D.J., Phillips, M.L., 2008. Elevated left and reduced right orbitomedial prefrontal fractional anisotropy in adults with bipolar disorder revealed by tract-based spatial statistics. Arch. Gen. Psychiatry 65, 1041–1052.

Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M., 2009. Bayesian analysis of neuroimaging data in FSL. Neuroimage 45, S173–S186.

Yakovlev, P.I., Lecours, A.R., 1967. The myelogenetic cycles of regional maturation of the brain. In: Minkowski, A. (Ed.), Regional Development of the Brain Early in Life. Blackwell Scientific Publications Inc, Oxford, England, pp. 3–70.

CHAPTER 5

Along-tract statistics

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Along-tract statistics allow for enhanced tractography analysis

John B. Colby [a,b,c], Lindsay Soderberg [c], Catherine Lebel [a,c], Ivo D. Dinov [a,d,e], Paul M. Thompson [a,b,d], Elizabeth R. Sowell [a,c,f,*,1,2]

[a] Department of Neurology, University of California Los Angeles (UCLA), Los Angeles, CA, USA
[b] UCLA Interdepartmental Program for Biomedical Engineering, Los Angeles, CA, USA
[c] Developmental Cognitive Neuroimaging Laboratory, Children's Hospital Los Angeles, Los Angeles, CA, USA
[d] UCLA Laboratory of Neuro Imaging (LONI), Los Angeles, CA, USA
[e] UCLA Department of Statistics, Los Angeles, CA, USA
[f] Department of Pediatrics, University of Southern California (USC), Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 24 June 2011
Revised 19 October 2011
Accepted 2 November 2011
Available online 9 November 2011

Keywords:
White matter
Tractography
Diffusion imaging
FASD
B-spline
Along-tract

ABSTRACT

Diffusion imaging tractography is a valuable tool for neuroscience researchers because it allows the generation of individualized virtual dissections of major white matter tracts in the human brain. It facilitates between-subject statistical analyses tailored to the specific anatomy of each participant. There is prominent variation in diffusion imaging metrics (e.g., fractional anisotropy, FA) *within* tracts, but most tractography studies use a "tract-averaged" approach to analysis by averaging the scalar values from the many streamline vertices in a tract dissection into a single point-spread estimate for each tract. Here we describe a complete workflow needed to conduct an *along*-tract analysis of white matter streamline tract groups. This consists of 1) A flexible MATLAB toolkit for generating along-tract data based on B-spline resampling and compilation of scalar data at different collections of vertices along the curving tract spines, and 2) Statistical analysis and rich data visualization by leveraging tools available through the R platform for statistical computing. We demonstrate the effectiveness of such an along-tract approach over the tract-averaged approach in an example analysis of 10 major white matter tracts in a single subject. We also show that these techniques easily extend to between-group analyses typically used in neuroscience applications, by conducting an along-tract analysis of differences in FA between 9 individuals with fetal alcohol spectrum disorders (FASDs) and 11 typically-developing controls. This analysis reveals localized differences between FASD and control groups that were not apparent using a tract-averaged method. Finally, to validate our approach and highlight the strength of this extensible software framework, we implement 2 other methods from the literature and leverage the existing workflow tools to conduct a comparison study.

Introduction

Since the late 1990s, diffusion magnetic resonance imaging (MRI) tractography methods have developed into a powerful set of techniques to investigate white matter connectivity in the human brain (Basser et al., 2000; Conturo et al., 1999; Jones et al., 1999; Le Bihan, 2003; Mori et al., 1999). By generating virtual dissections of different white matter tracts for each individual, tractography has proved valuable in a variety of applications — including pre- and intra-operative mapping of fiber tracts (Duncan, 2010; Maruyama et al., 2005; Prabhu et al., 2011; Young et al., 2010), and connectivity analyses of anatomical and functional brain networks (Aron et al., 2007; Behrens et al., 2003; Bullmore and Sporns, 2009; Ramnani et al., 2004). In addition to providing information about tract geometry, tractography can provide individualized volumes of interest for the investigation of white matter microstructural *qualities* in the context of development and aging (Asato et al., 2010; Davis et al., 2009; Eluvathingal et al., 2007; Huang et al., 2006, Lebel et al., 2008b, 2010; Liston et al., 2006; Penke et al., 2010; Sala et al., in press; Schmithorst and Yuan, 2010; Verhoeven et al., 2010; Voineskos et al., 2012), numerous diseases (Ashtari et al., 2007; Kumar et al., 2010; Kunimatsu et al., 2003, Lebel et al., 2008a; Zarei et al., 2009), and the relation of brain structure to functional, cognitive, and psychiatric differences between individuals (Boorman et al., 2007; Dougherty et al., 2007; Glenn et al., 2007; Lebel and Beaulieu, 2009; Luck et al., 2011; Schulte et al., 2010; Tsang et al., 2009). As tract dissections are personalized to each individual, and do not rely on any between-subject warping to a common template space, analogous regions can be compared between individuals even when there are large differences in brain morphology. This is valuable in

* Corresponding author at: Children's Hospital Los Angeles, 4650 Sunset Blvd., Mailstop #130, Los Angeles, CA 90027, USA. Fax: +1 323 361 7836.
E-mail address: esowell@chla.usc.edu (E.R. Sowell).
[1] Director, Developmental Cognitive Neuroimaging Laboratory.
[2] Professor of Pediatrics, Keck School of Medicine, University of Southern California.

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.11.004

clinical studies, in which patients might have gross structural brain abnormalities through alterations in neurodevelopment, or in white matter microstructure as a result of disease.

Direct voxelwise comparison of diffusion imaging data is challenging, as the high-contrast edges of diffusion imaging volumes (e.g., FA maps) make them more susceptible to small misregistration errors, as well as to anatomical variability of tract position in health and disease. Even so, traditional voxelwise brain mapping is an important complement to tractography — especially now that analysis methods have advanced beyond a generic voxel-based statistical approach to include more optimized strategies tuned specifically for the analysis of white matter and diffusion imaging data (Smith et al., 2006, 2007). Furthermore, the inherent voxel-to-voxel independence of voxelwise processing allows these methods to see beyond the type of small focal disruptions that could potentially derail the streamline tractography algorithms. In general, these voxelwise studies have been in broad agreement with their tract-based counterparts (Schmithorst and Yuan, 2010; Sullivan and Pfefferbaum, 2006; Wozniak and Lim, 2006). Additionally, they have demonstrated a remarkable degree of regional heterogeneity – even *within* a given tract – in the diffusion imaging indices and observed relationships with other variables (Barnea-Goraly et al., 2010; Bava et al., 2010; Bengtsson et al., 2005; Hsu et al., 2010; Keller and Just, 2009).

To improve the localizability in deterministic tractography, there is a growing interest in methods that can provide greater *within*-tract detail. While previous work has included efforts to examine DTI metrics along tract lengths (Corouge et al., 2006; Goodlett et al., 2008, 2009; Jones et al., 2005; O'Donnell et al., 2009; Zhu et al., 2010), as well as more generic within-tract methods that can accommodate variability along even more dimensions within tracts (Yushkevich et al., 2008; Zhang et al., 2010), it has typically been focused on individual aspects of the along-tract workflow or specific customized applications. It remains true that the large majority of streamline tractography analyses still rely on a simpler tract-averaged methodology. Therefore, there is a need for a higher-level integrated along-tract processing workflow — for an intuitive set of tools that makes it easy for applications researchers to start incorporating along-tract detail into existing tractography analyses, while facilitating statistical analysis of these data in a general linear model (GLM) framework, visualization of raw data and statistical results, and straightforward customization of all aspects of this process. To help fill this gap in the methods landscape, in this manuscript we: 1) Describe the rationale for conducting a tractography study with enhanced within-tract detail, as it relates to common tractography applications within the neuroimaging community, 2) Lay out a straightforward workflow for conducting one type of *along*-tract analysis, which is able to attain a useful balance between accessibility and improved modeling ability, 3) Demonstrate some advantages of this approach over traditional tract-averaged methods by looking at both within-subject and between-group examples, 4) Validate this approach against existing methods, while highlighting the extensible nature of this workflow toolset, and 5) Make this generic toolset available for others to use as building blocks for their own future analyses (http://www.github.com/johncolby/along-tract-stats).

Along-tract statistics

Rationale

When standard tractography methods collapse tract groups, they yield only a single mean DTI metric and variance estimate for each tract and for each subject. This processing step ignores the potentially rich anatomical variation in diffusion imaging metrics *along* the tracts, and reduces the effectiveness of this technique. To see that this added detail exists, one can browse through an FA map, or look at a histogram

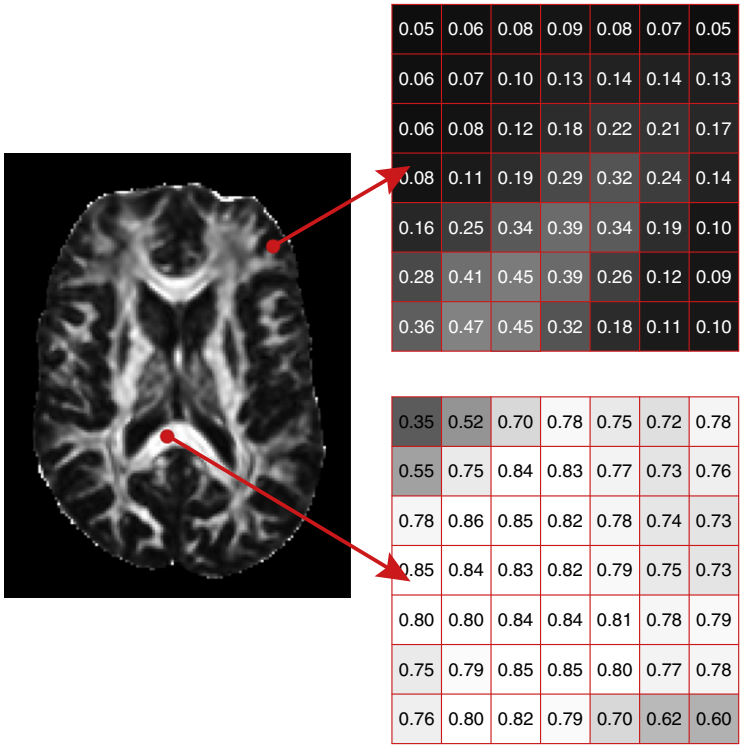


Fig. 1. FA variations throughout the brain: Fractional anisotropy (FA) varies widely throughout the white matter, with values ranging from below 0.2 at the transition to gray matter near the cortex (top breakout panel), to greater than 0.8 in tightly coherent fiber bundles like the midline corpus callosum (bottom breakout panel).

of its contents. FA varies widely throughout the white matter — with very low values (below 0.2) at the transition to the gray matter of the cortex, and very high values (above 0.8) in highly coherent areas like the midline corpus callosum (Fig. 1). However, when the tract-averaged estimates of mean FA in the major white matter tracts are examined, they generally fall between 0.4 and 0.6 (with the majority between the even tighter range of 0.4 to 0.5) (Wakana et al., 2007). This suggests that a large amount of blurring is taking place in the data and that potentially interesting features are being lost. Another useful way to see the extent of within-tract variability is to overlay the scalar measure of interest (e.g., FA) onto the streamlines themselves (Fig. 2). Here again it is easy to see the amount of detail within tracts. This representation, in particular, begs the analogy to a highway system: There is a collection of roads (tracts) with different speeds (FA). While there is some variation in average speed between highways (some highways are *always* slower than others), and some interesting inferences can even be made this way (traffic might universally be worse on the way to work than at midnight), to get the most complete

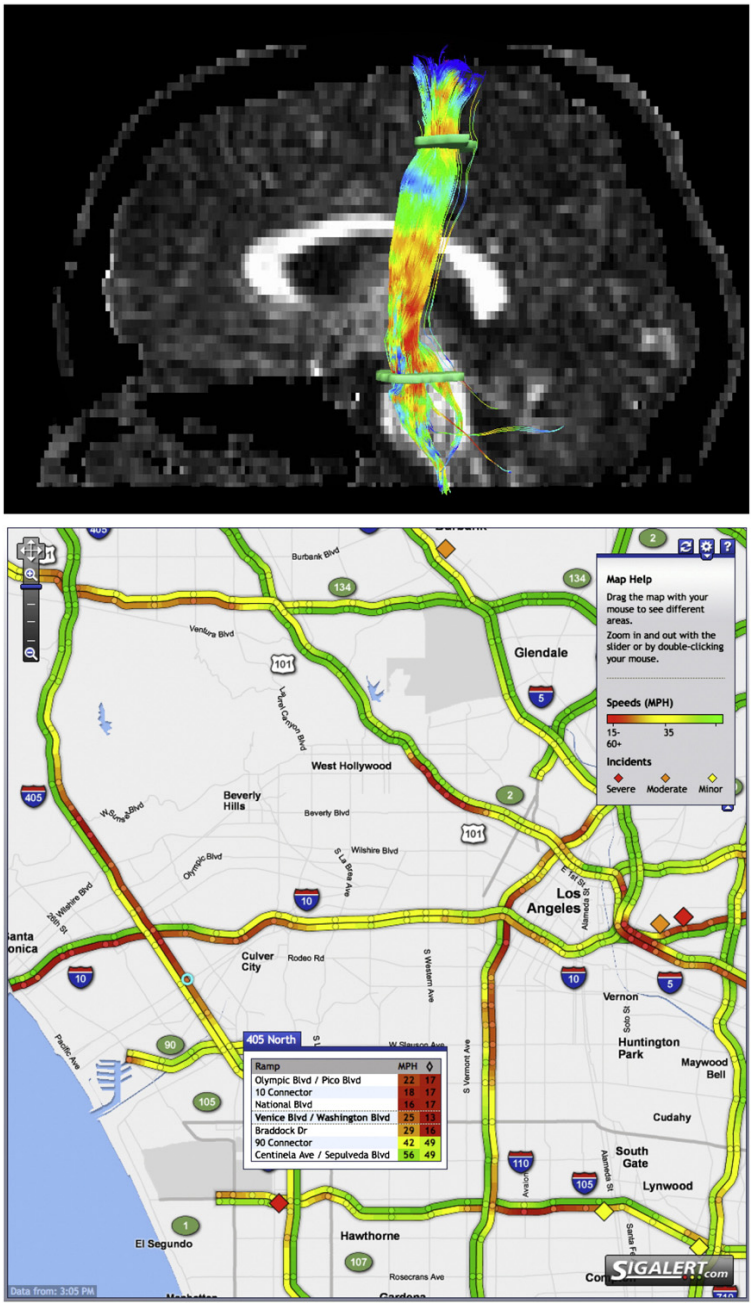


Fig. 2. Along-tract variations in the human brain and the Los Angeles highway system: (Top panel) Deterministic tractography dissection of the left corticospinal tract in one individual. Color is used to encode FA variations along the tract. (Bottom panel) Highway map of Los Angeles, CA. Color is used to encode traffic speed.

picture of what is happening in the system, one must consider the *within*-highway effects that can have large influences (e.g., car crashes, construction, and lane geometry can all have profound focal effects on the flow of traffic). Shifting back to the brain, there is a similarly strong neurobiological basis for within-tract variability, as the vascular (Ishii et al., 1996; Ito et al., 2005), supporting glial (Innocenti et al., 1983; Monier et al., 2006, 2007; Yeh et al., 2009), biochemical (McIntosh et al., 2008; Moghaddam and Adams, 1987; Pankonin et al., 2009; Perry et al., 1971; Stamford et al., 1984; Vasung et al., 2010; Warrington et al., 2007), and biophysical environments (Childs et al., 2007) are all found to vary throughout the interior of the brain.

Beyond true *biological* variation in white matter properties along tracts, the local disturbance of diffusion imaging parameters due to *methodological* issues is another reason that makes it attractive to analyze these properties along tracts. For example, complex fiber geometries like crossing, kissing, and partial volume averaging in general can alter properties like FA while more direct measurements of myelin content may remain unchanged (Stikov et al., 2011). Further rationale comes from the difficulty in corroborating results across studies. For any given diffusion imaging application, it is common to find that some studies have used voxelwise methods like tract-based spatial statistics (TBSS) (Smith et al., 2007), while a different collection may have used tract-averaged streamline tractography. Comparing the two types of studies is not always straightforward, and can lead to puzzling questions, like: Where along the tract-averaged ROI is the affected region located? Is it a constant difference along the whole tract, or a more focal abnormality? How does a voxelwise cluster in the internal capsule in one study relate to a significant tract-averaged tractography finding across the corticospinal tract in another? Are the same voxels even included in both studies? Attaining a higher degree of within-tract detail in our tractography analyses may help improve the level of interoperability and comparability between tractography and voxelwise methods.

Importantly, this general approach is not meant to replace the standard tract-averaged analysis techniques, but to provide the option for additional detail. Conversely, this report is not intended to be a focused exploration of the lower-level nuances between the different resampling, point alignment, and other mathematical parameters involved. Rather, we will emphasize a high-level perspective on the broader practical considerations of applications scientists, and intend to convince the reader that by making these minor modifications to existing tractography workflows, one can map the along-tract detail in the brain and enhance a broad range of white matter tractography analyses.

Overview

We aimed to create an intuitive and flexible set of modular tools with a balance between within-tract modeling complexity and accessibility. We tapped into established computational, statistical, and visualization libraries where possible, and, where decisions on low level processing approaches or parameters were needed, we attempted to make rational "middle-of-the-road" choices. Again, the overall goal was to create an end-to-end workflow that would be a useful and flexible starting point for applications-oriented neuroscientists who would otherwise conduct a standard tract-averaged tractography analysis. In short, for a given tract group, we 1) Reorient the streamlines according to a common origin, 2) Re-parameterize the streamlines with cubic B-splines, 3) Resample the streamlines so that each has the same number of points spread evenly along its length, 4) Resample the underlying voxel volume at these new vertices, and then 5) Collapse these values across streamlines at each analogous group of vertices to obtain mean scalar estimates at many locations along the tract. This allows FA or other scalar metrics to be analyzed between subjects/groups at each collection of vertices along the length of the tract, instead of using one overall tract-averaged value.

The included toolset (http://github.com/johncolby/along-tract-stats) allows the reader to begin with their own raw streamline tract groups and proceed through all the steps necessary to perform a complete along-tract analysis (these steps are described next in Processing workflow section). These tools may be operated one by one in an exploratory interactive manner, or automated in a batch mode to streamline an entire between group analysis through to the creation of customized statistical plots like those seen in this manuscript. Available to the reader are full documentation of all processing tools, an online interactive web demonstration of along-tract methods (http://www.openprocessing.org/visuals/?visualID=25715), and online tutorials including example data to download and full example analyses (http://github.com/johncolby/along-tract-stats/wiki).

Processing workflow

Preprocessing

Before calculating along-tract statistics, the tract groups must be delineated. There are many available software platforms to do this. We used tract groups delineated manually for each subject in TrackVis (http://www.trackvis.org). Because these tools are modular functions written in plain text in MATLAB, this framework can be straightforwardly adapted to operate on streamline data from a variety of sources. First, a tensor or alternative diffusion model is fit to the raw data, and the resulting directionality information at each voxel is used to generate a collection of streamlines representing white matter pathways across the entire brain. This initial set of fibers must then be pared down to just the groups that comprise each tract. The means of achieving this virtual dissection can vary, and a variety of atlas-based (Lebel et al., 2008b) or more unsupervised clustering (Clayden et al., 2007; Maddah et al., 2008) approaches can be attractive options depending on the size of the dataset and other considerations. Nevertheless, for between-group comparisons with clinical populations, the gold standard remains manual ROI-based dissection by a trained experimenter according to a reliable protocol. While the following descriptions will focus on FA maps from diffusion tensor imaging (DTI) data, these techniques are more generic in nature and could be applicable in a variety of other scenarios (e.g., vertex output of other packages, higher order diffusion models, multi-streamline/probabilistic tractography algorithms, mean diffusivity maps, etc.).

Reorient streamlines

Streamlines generated using brute force tractography algorithms (such as the standard Fiber Assignment by Continuous Tracking (FACT) algorithm (Mori et al., 1999) used in TrackVis) are naïve to any sort of tract origin. For example, when streamlines are generated that will end up comprising the corticospinal tract, there is no logic to say whether the streamlines are ordered to "start" at the cortex end or the brainstem end of the tract. To proceed with the cross-sectional along-tract modeling, the streamlines in each tract group must be reoriented so that they all "start" at the same end. For some relatively linear tracts, like the corticospinal tract, this can be automated. For example, the starting end of corticospinal tract streamlines could always be chosen to be the end closest to the lowest axial slice. For other tracts, like the arcuate fasciculus, however, intersubject variation in tract position and in-scanner positioning makes this less straightforward. In these cases, some sort of in-line interactive assignment may make the most sense. For example, a tract group can "pop up" on the screen and the user can click on the end that they want to designate as the "origin". See Table 1 for a list of the tract origin conventions used in this report.

Table 1
Single subject tract information and ANOVA results between along-tract and tract-averaged models: Hemisphere, tract origin, number of streamlines, and number of resampled vertices are tabulated for each tract group in the single subject atlas (Fig. 5). The along-tract and tract-averaged linear models were fit to these streamline data, and ANOVA was used to determine if moving to the along-tract approach provided significant increases in explained variance over the tract-averaged form.

Tract	Hemisphere	Origin	Streamlines	Resampled vertices	F	df (numerator, denominator)	*p*-value
Cingulum–cingulate gyrus part	L	Anterior	73	30	48.4	29, 2160	<2.2e-16
Cingulum–hippocampal part	L	Anterior	71	14	12.4	13, 980	<2.2e-16
Corticospinal tract	L	Inferior	180	47	138.2	46, 8413	<2.2e-16
Anterior thalamic radiations	L	Anterior	123	31	43.3	30, 3782	<2.2e-16
Arcuate fasciculus	L	Frontal	125	40	72.8	39, 4960	<2.2e-16
Inferior longitudinal fasciculus	L	Anterior	241	41	216.0	40, 9840	<2.2e-16
Inferior fronto-occipital fasciculus	L	Anterior	99	62	93.6	61, 6076	<2.2e-16
Uncinate fasciculus	L	Frontal	89	28	48.0	27, 2464	<2.2e-16
Corpus callosum — forceps major	–	Right	94	64	142.8	63, 5952	<2.2e-16
Corpus callosum — forceps minor	–	Right	496	41	326.0	40, 20295	<2.2e-16

Model streamlines with cubic B-splines

The main hurdle to conducting an *along*-tract analysis at many locations within a tract is that the number of vertices that make up a streamline can vary between streamlines and subjects. Further, the spacing of the vertices within each streamline is also variable, resulting from the way the vertices are laid down as the streamlines traverse the underlying cuboidal voxel structure. One natural option, which we use to address this issue here, is to use cubic B-splines to re-parameterize the polylines as curves (a schematic is shown in Figs. 3A,B) (Corouge et al., 2006; O'Donnell et al., 2009; Zhu et al., 2010). Fig. 3 is also useful to demonstrate that by "cross-section", we mean a collection of analogous points across streamlines, rather than a simple two-dimensional slice plane.

Resample streamlines with constant number of vertices

Once the streamlines have been approximated using cubic B-splines, they can be resampled according to different criteria. Perhaps the most straightforward option, here we focus on an approach that resamples each streamline into the same number of vertices spread evenly along their lengths, regardless of the length of the streamlines (Figs. 3C,D). This type of method accounts for inter-streamline and inter-subject scaling, and facilitates between-subject statistical

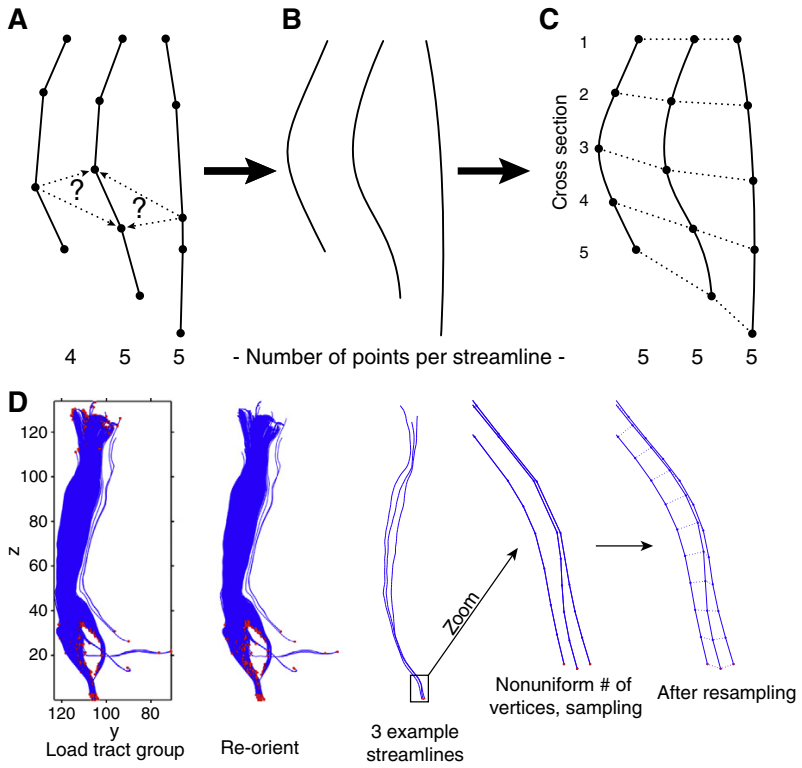


Fig. 3. B-spline based resampling: (A) Raw deterministic tractography streamlines are difficult to analyze along their length due to differing numbers of vertices, and non-uniform spatial sampling. To address this, the raw streamlines can be re-parameterized using cubic B-spline curves (B), and then resampled to allow for a straightforward analysis at different tract "cross-sections" (C). (D) Streamline processing for an example tract group (left corticospinal tract) derived from actual data. The streamlines are first reoriented so that their origins (red points) are near a common tract terminus, and then resampled to allow for comparison across streamlines at different cross-sections (dotted lines, right).

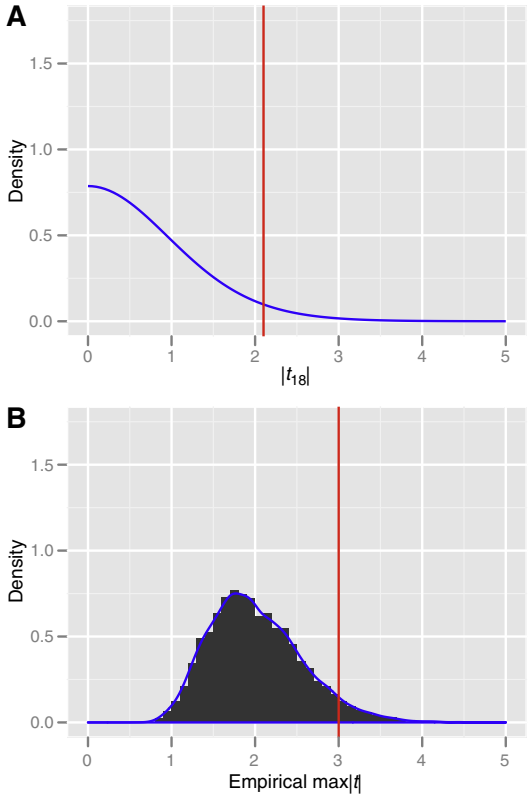


Fig. 4. Correcting for multiple comparisons: (A) Theoretical probability density function and $p<0.05$ critical value for a single two-tailed *t*-test with 18 degrees of freedom. (B) The empirical null distribution of the *maximum* test statistic across the many along-tract comparisons from an example analysis. The $p<0.05$ critical value has been shifted appropriately to control the family-wise error rate (i.e. the chance of *any* false positives across *all* of the multiple comparisons).

analysis by maintaining a one-to-one mapping of the vertices at each cross-section, and across streamlines and subjects.

Extract scalars and collapse cross-sectionally

The final step in compiling along-tract data is to resample the underlying scalar volume (e.g., FA map) at the new sets of streamline vertices. For each subject, these scalars are then averaged at each analogous set of vertices along the tract, to obtain cross-sectional scalar mean and variance estimates. An important note with this approach is that each vertex contributes to these mean and variance estimates. Because there is usually a higher streamline density toward the center of a tract, these values will typically be weighted to favor the cores of tracts where multiple streamlines pass through the same voxels. While this is something that may or may not be desirable, given a particular application, it is an important and interesting topic to consider when performing along-tract analysis. The spatial location of the vertices can also be averaged to obtain a single streamline that represents the mean tract geometry. This aids visualization by displaying the along-tract scalar estimates of a single subject, and statistical results of between-group analyses (see Figs. 5 and 7). Again we stress that because of the flexibility this implementation allows, our processing decisions tend towards straightforward options like this in order to allow operators to develop an intuitive understanding of the broader workflow while still facilitating their implementation of other variations on these techniques if so desired at a later time (For example, see Toolkit extensibility section for an example implementation of several other options from the literature, and O'Donnell et al. (2009) for a discussion of the variability to expect among these approaches).

Between-group statistical analysis

Linear mixed-effects model

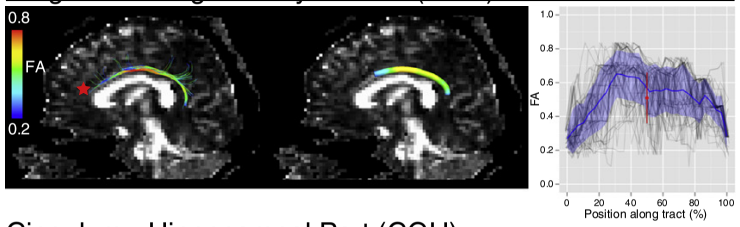
With cross-sectional mean FA and variance estimates obtained at many locations along a tract, and at analogous anatomical locations in each individual, we now fit statistical models to these data. This statistical analysis and later data visualization are performed using R (http://www.r-project.org), a flexible platform that is emerging as a powerful source for neuroimaging analysis resources (Tabelow et al., 2011). A linear mixed-effects (LME) model is applied serially for each tract group (i.e. for each tract and hemisphere) (Pinheiro and Bates, 2000). Fixed effects include: an overall intercept, a "position" factor (dummy coding the tract cross-section indices as levels of an unordered factor), a "group" factor (i.e. Control or Patient), and a "group:position" interaction. A subject level random effect term is also included to explain the variance component associated with this repeated measures design. Standard analysis of variance (ANOVA) can then be used to test the significance of these sets of terms in the model. The intercept term tests whether the overall grand mean FA is different from zero, which will be true by definition since there will be no streamlines if there is no FA. The group term tests whether there is an overall effect of group membership on the FA vs. position curve (analogous to the traditional tractography analysis that tests for changes in tract-averaged mean FA). While this example uses the case of a binary covariate testing for a group difference, the general linear model (GLM) can also accommodate multilevel factors or continuous covariates. The F-test across the position terms tests whether FA is, in fact, varying within the tract. Finally, the F-test across the group:position interaction term determines whether the effect of group on FA varies based on position. In other words, it tests for any more *focal* regions of group effects isolated to regions *along* the tract. Another option would be to use a multivariate approach; either option is valid, and the choice depends on the types of questions researchers want to ask.

Multiple comparisons

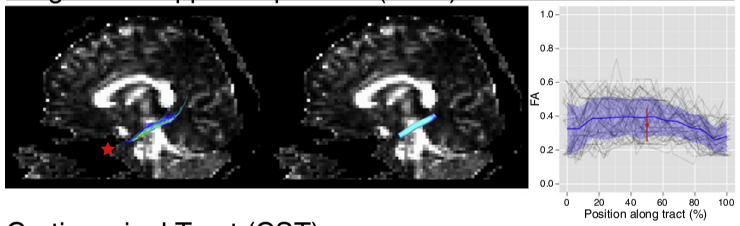
One consequence of choosing the serial univariate approach is that the family-wise Type 1 error rate of the individual t-tests along the tracts, if examined directly, will be inflated due to the increased number of multiple comparisons (Shaffer, 1995). To address this, we can apply permutation methods to adjust the *p*-values and control the Type 1 error. Using the test for group differences as an example, we assume there is no group effect under the null hypothesis, and therefore that the group labels are exchangeable. The group labels can then be permuted, the model fit once again, and the maximum statistic across *all* comparisons recorded. This process is repeated many times to empirically build up the distribution of the *maximum* test statistic under the null hypothesis. To determine how extreme the test statistics obtained from the model fit to the real data are,

Fig. 5. Single subject along-tract analysis: The raw streamlines (left panels), mean tract geometries (middle panels), and along-tract variations in FA (right panels) are displayed for 10 major white matter tracts in the human brain. In the two streamline views, color is used to encode variation in FA. In the along-tract plots, FA is plotted versus position from tract origin (designated as the side of the tract near the red star in the streamline views; see also Table 1). The distribution of individual streamlines is shown in the background (black lines; transparency and slight x-axis jitter used to control overplotting). Overlaid is the along-tract cross-sectional mean FA (blue; ± pointwise standard deviation). Also included is the standard tract-averaged point-spread estimate (red).

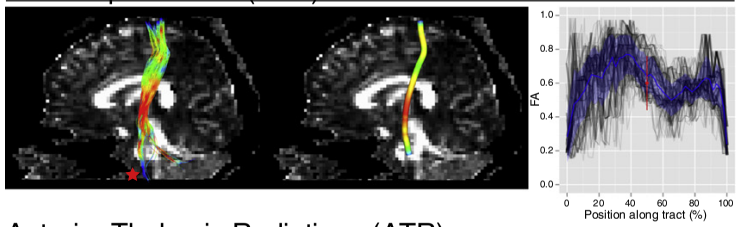
Cingulum - Cingulate Gyrus Part (CGC)



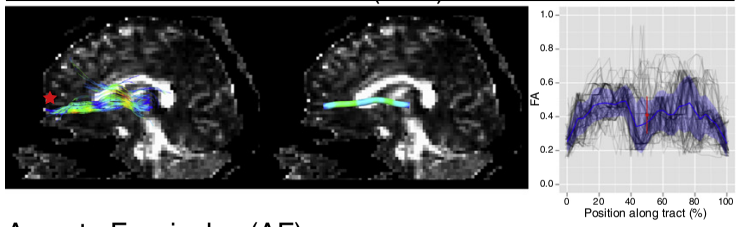
Cingulum - Hippocampal Part (CGH)



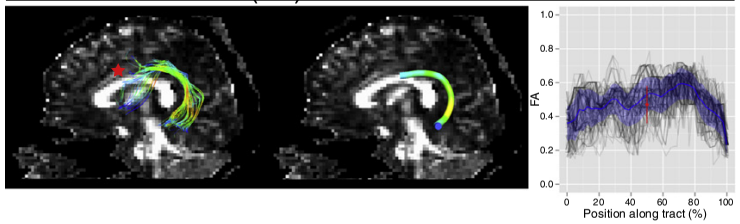
Corticospinal Tract (CST)



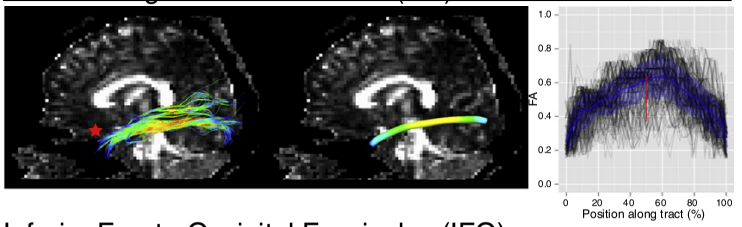
Anterior Thalamic Radiations (ATR)



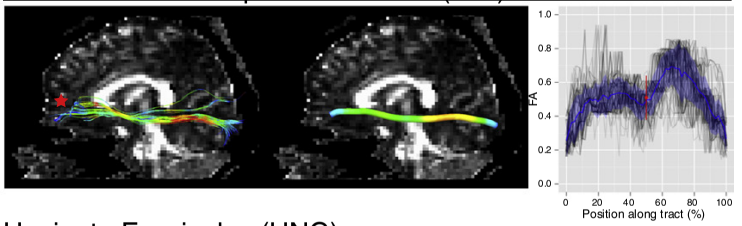
Arcuate Fasciculus (AF)



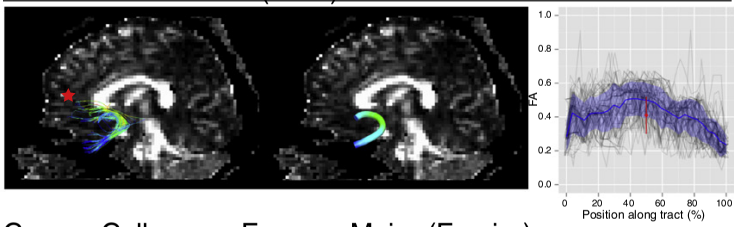
Inferior Longitudinal Fasciculus (ILF)



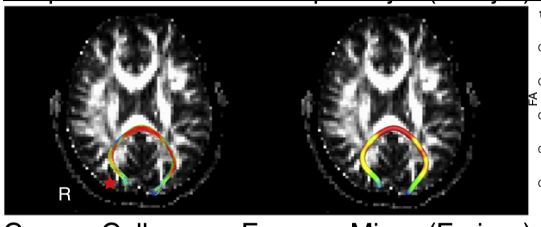
Inferior Fronto-Occipital Fasciculus (IFO)

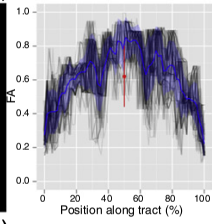


Uncinate Fasciculus (UNC)

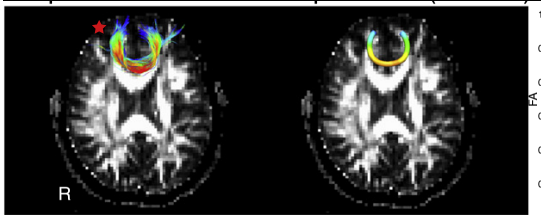


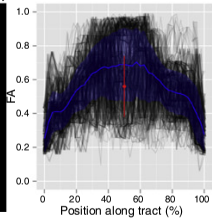
Corpus Callosum - Forceps Major (Fmajor)





Corpus Callosum - Forceps Minor (Fminor)





one can simply compare them to this null distribution in order to obtain *p*-values corrected for multiple comparisons (Nichols and Hayasaka, 2003; Nichols and Holmes, 2002). This same procedure can also be applied to correct the *p*-values associated with other covariates of interest (age, cognitive measures, etc.) (Fig. 4).

Visualization

Two dimensional (see Fig. 6)

Along-tract methods increase the amount of data that results from a tractography analysis (for example, Fig. 10 contains on the order of 200,000 data points), so it is important to consider how this expanded data structure is presented to the viewer in order to maximize the usefulness of the data for various tasks. We begin with a single panel, which displays the scalar metric plotted as a function of position along the tract, and colored according to group membership. The reader is given access to the raw data for each subject, displayed semi-transparently in the background, as well as a higher-level statistical summary like the smoothed estimate of the group mean $\pm$ 95% pointwise confidence interval. Annotations can also be overlaid to convey the results of hypothesis testing — for example, an asterisk in the corner if there is an overall offset between groups, or, if the group:position F-statistic is significant, a bar to signify which component statistical tests are significant. The natural divisions in the data by tract and hemisphere can then be used to generate a facetted display of these panels, which allows for efficient review of the *entire* dataset once the reader becomes familiar with a single panel (Tufte, 2001). The panels are placed in close context, so the researcher can quickly review differences between hemispheres and tracts, and explore multivariate patterns *across* tracts and hemispheres. Whole tract properties, like the number of streamlines, may also be displayed in adjacent panels of bar or box plots. These data can even be encoded into the main panels through other plotting esthetics such as line width. This could be used to draw attention to individuals with many streamlines or to check if outliers have too few. Finally, it is important that the generation of these statistical graphics be automated and reproducible, to provide robustness to operator errors, and easy extensibility/portability to new datasets and applications. To achieve this, we employ an implementation of the "grammar of graphics" in R (http://had.co.nz/ggplot2), which provides a rich set of abstracted graphical language tools to generate these complex multilayered plots (Wickham, 2009; Wilkinson, 2005).

Three dimensional (see Fig. 7)

As a complement to the 2D statistical graphics, it is also useful to view the statistical results overlaid on the tract geometry. To do this, the mean tract geometries for all of the tract groups for a representative subject can be displayed together. However, in addition to using color to encode FA or direction, we can use it to visualize the effect sizes and *p*-values from a statistical analysis. This may be particularly useful when comparing tractography results to voxelwise results – either from the same individuals or from other analyses reported in the literature – as it directly connects these two methods and makes it easier to corroborate results.

Example implementation

Data acquisition and preprocessing protocol

For the within-subject and between-group analyses below (Within-subject analysis and Between-group analysis sections), whole brain diffusion weighted imaging data were acquired on a 3 T Siemens Trio MRI scanner. Each DTI acquisition included diffusion weighted volumes (30 directions, $b = 1000$ s/mm^2, 240 mm field of view, 96×96 in-plane matrix, 55 axial slices of 2.5 mm thickness, resulting in $2.5 \times 2.5 \times 2.5$ mm^3 isotropic voxels), and one non-diffusion-weighted volume ($b = 0$ s/mm^2). A tensor model of diffusion was fit to these raw data, and scalar maps of FA were generated. Whole-brain brute force deterministic tractography was performed according to the FACT algorithm (Mori et al., 1999), as implemented in Diffusion Toolkit v0.6 (http://www.trackvis.org/dtk). Tracking constraints included a minimum FA threshold of 0.15 and a maximum fiber turning angle of 60°. Tract groups were then manually extracted in TrackVis v0.5.1 (http://www.trackvis.org) by a trained experimenter (L.S.) according to the instructions from Wakana et al. (2007). Human subjects data were collected as part of an ongoing study that has been approved by the institutional review board at UCLA.

Within-subject analysis

To validate this along-tract approach relative to the traditional tract-averaged method, we compared the two analyses to each other for ten tracts in one healthy young adult subject. The resulting tracts were compared 1) visually, by using the streamlines obtained from manual delineation and the mean tract overlaid with average cross-sectional FA values (see Fig. 5), and 2) quantitatively, by including data from all streamlines for each tract in a simple linear model, and performing an ANOVA between the along tract model (FA as a function of an intercept term and position terms) and the tract-averaged model (FA as a function of an intercept term only). As expected, along-tract processing provided highly significant increases in the explained variance for all tracts (see Table 1).

Between-group analysis

To demonstrate how easily this approach extends to between-group analyses, we conducted a comparison between a group of children with fetal alcohol spectrum disorders (FASDs) ($n = 9$, age $= 13.8 \pm 2.6$ years, 4 females) and typically developing controls ($n = 11$, age $= 13.2 \pm 3.1$ years, 5 females). FA was analyzed along tracts bilaterally within the inferior longitudinal fasciculus (ILF) and the arcuate fasciculus (AF). The ILF was chosen because it is linear, fairly "rope-like", robustly trackable, and exhibits some of the largest group differences between these populations (Lebel et al., 2008a). Conversely, the AF was chosen as a more challenging example because it has a nonlinear geometry and is more difficult to track. The numbers of streamlines that made up these tracts were first analyzed between groups, revealing that 1) FASD subjects had significantly fewer streamlines than controls ($t = -2.35$, $df = 76$, $p = 0.022$), and 2) The ILF has significantly more streamlines than the AF ($t = 3.70$, $df = 76$, $p = 4.10 \times 10^{-4}$). Next, as might be typical of many neuroscience applications where the experimenter seeks to map the white matter tract abnormalities associated with a certain disease, an along-tract statistical analysis was set up to look for overall offsets in the FA vs. position curves between groups, and also to look for more localized regions of effects along the tracts (see Linear mixed-effects model section for a complete description). This protocol failed to reveal any overall offsets between groups, but along-tract analysis did demonstrate a significant group:position interaction in the left inferior longitudinal fasciculus (L ILF) ($F_{29,522} = 2.53$, $p = 2.64 \times 10^{-5}$). This effect localized to a region in the posterior portion of the tract where the FASD group had significantly lower FA than the control group ($t = -3.16$, $df = 521$, $p = 0.0017$, $p = 0.032$ corrected) (see bottom left panel in Fig. 6, and visualization in Fig. 7).

Toolkit extensibility

To demonstrate the strength of our toolkit's extensibility and modularity, and to further validate our approach, for comparison we implemented 2 other streamline correspondence schemes from the literature. The "Distance Map" method pre-computes the minimum

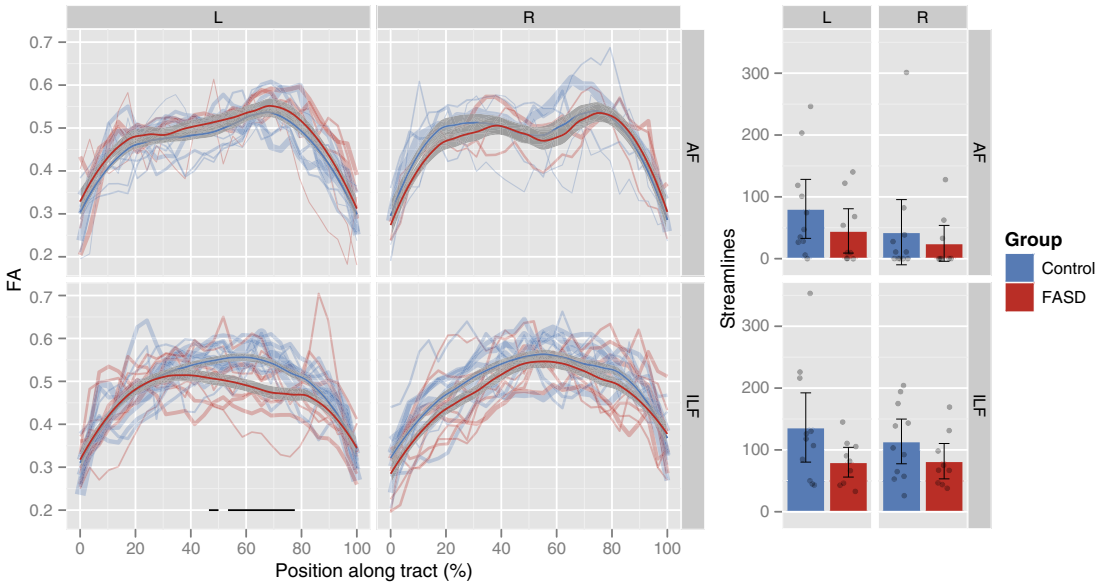


Fig. 6. Between-group along-tract analysis in fetal alcohol exposure: FA is plotted versus position from tract origin, with plots faceted by tract name and hemisphere, and colored according to group membership (Control or FASD). Along-tract estimates from individual subjects are displayed semi-transparently in the background (line width encodes the relative number of streamlines), and are overlaid with locally weighted smooth estimates of the group means (± pointwise 95% confidence range). The number of streamlines in each tract (± pointwise 95% confidence interval) is also included in an accompanying panel at right.

Euclidean distances from all points on a voxel-like grid to the vertices of a prototype fiber geometry, and uses this information to generate a lookup table describing which regions of the grid will be mapped to which regions of the prototype (Fig. 8 shows a 2-dimensional example). If multiple fibers share the same prototype (i.e., they are part of the same tract group, or different subjects' tract groups that have been registered to the same space), this has the advantage of dramatically speeding up the processing (Maddah et al., 2008). In a related "Optimal Point" matching approach, only the components of the Euclidean distances tangent to the prototype fiber geometry are considered, and matches are assigned via global cost optimization. This has the effect of strongly favoring matches that lie orthogonal to the prototype fiber (O'Donnell et al., 2009).

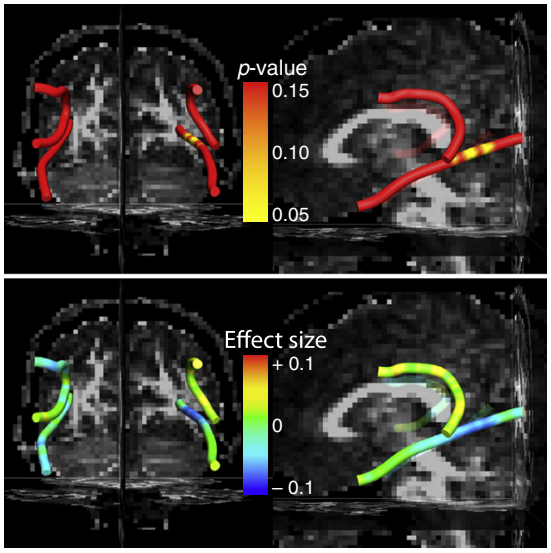


Fig. 7. Visualization of between-group results: Statistical results are displayed on the mean tract geometry of one representative subject. (A) *p*-value map (yellow color indicates regions of significant effects ($p<0.05$, corrected)). (B) Effect size map (cooler colors represent regions of decreased FA in FASD subjects).

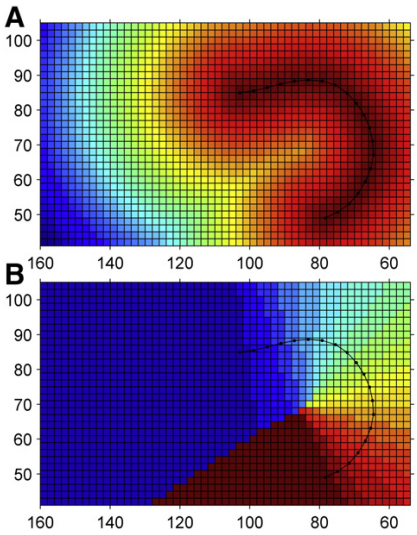


Fig. 8. Distance map method (2-dimensional): Euclidean distances are calculated from each point on a grid to each point on the prototype fiber, and used to generate a minimum distance map (A) and corresponding label map (B) matching regions of the grid to different points on the prototype. This lookup table can then be used for rapid processing of multiple fibers that share the same prototype (Maddah et al., 2008).

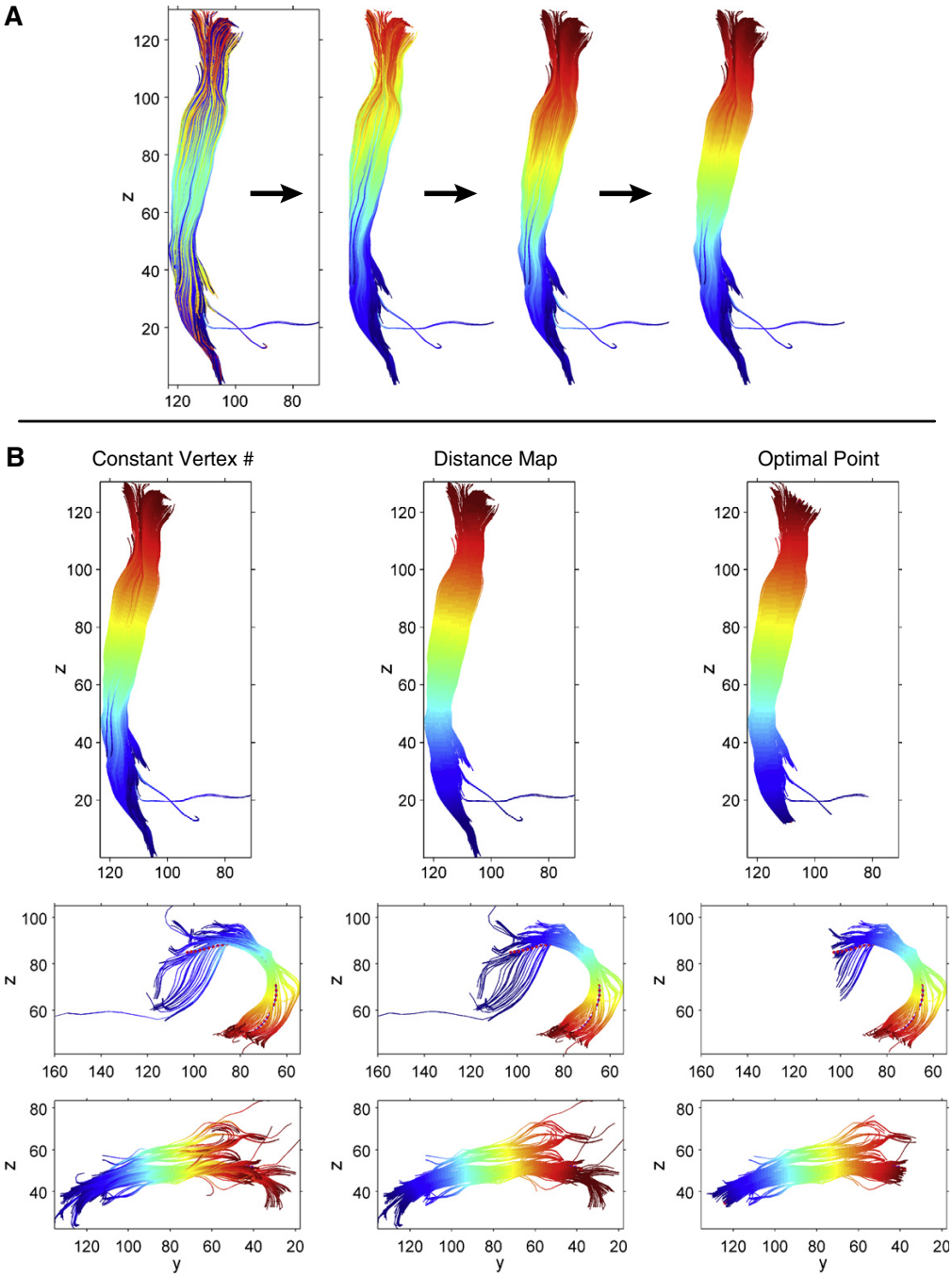


Fig. 9. Correspondence plots: Streamlines are colored by vertex index, highlighting which vertices have assigned correspondence and will be grouped together for the analysis (i.e. all vertices with the same dark blue hue will be grouped together, etc.). (A) Left corticospinal tract, after 1) import of raw streamlines, 2) reorientation of streamlines toward a common origin, 3) resampling of streamlines to have the same number of vertices, and 4) automatic prescription of an additional interior point of correspondence. (B) Comparison of 3 different correspondence schemes (columns; Constant Vertex #, Distance Map, Optimal Point), for 3 different tract files (rows; left corticospinal tract, left arcuate fasciculus, left inferior longitudinal fasciculus). The mean tract geometry (i.e., the prototype fiber) is also plotted, and is visible where not obscured by the other fibers.

A plug-in was generated with these algorithms, and their alternate correspondence labels were easily incorporated into our existing data structure. Leveraging our tools already available to the user, facetted multi-panel figures of streamline correspondence plots (Fig. 9) and along-tract data plots (Fig. 10) were then generated to allow for comparison of these different methods for 3 different example tracts.

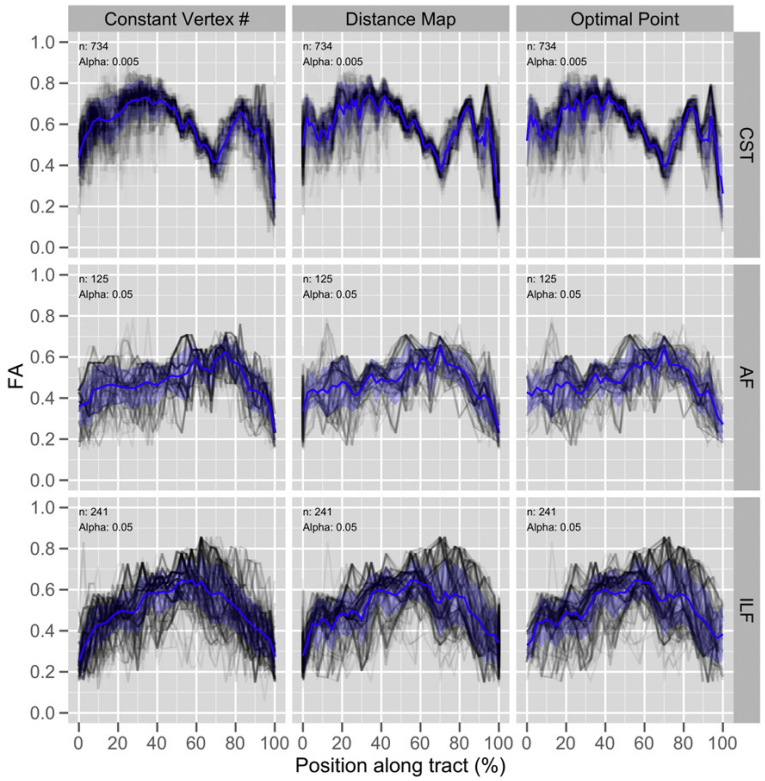


Fig. 10. Along-tract plots using different correspondence schemes: FA is plotted versus position from tract origin (similar to Fig. 5). The distribution of individual streamlines is shown in the background (black lines; transparency and slight x-axis jitter used to control overplotting). Overlaid is the along-tract cross-sectional mean FA (blue; ± pointwise standard deviation). Plots are annotated with number of streamlines (n), and transparency value used (alpha). As in Fig. 9, these plots are facetted into a 3×3 grid of different correspondence schemes (columns), and tracts (rows).

Discussion

Processing workflow

The main goal of this work was to generate a simple, flexible end-to-end workflow for along-tract processing and statistical analysis. We focused on the portion of the within-tract variability that exists along tracts because: 1) Previous studies show that the largest component of the within-tract variability exists along this axis, and 2) Tract groups already have longitudinal structure built in, due to the connectivity of adjacent vertices in each streamline. Clearly this decision is more appropriate for some tract geometries than others, and suggests that these methods are best suited for tract dissections that are relatively long (there is less benefit to analyzing differences along a tract if the tract is so short that there is little change along its length) and restricted to relatively tube-like point-to-point trajectories between functional or anatomical regions. For example, an analysis of the corticospinal tract's projections from the primary motor cortex would be more appropriate than an analysis that includes the fanning geometry of the *entire* corticospinal tract. Whether the assumption of radial uniformity is valid is also somewhat operator dependent, and will vary based on the types of hypotheses and effect sizes expected in the data.

We used a "constant number of vertices" approach to resample streamlines because it is simple; it implicitly scales all streamlines to the same length, controlling for variation in tract length between streamlines and subjects, and providing a natural tract "origin" at either end of the tract; and it allows straightforward between-group analysis because every subject's tract data will have a one-to-one mapping along the curved spine of this "position" axis. Although the optimal number of resampling points is expected to vary by many factors including spatial resolution, the smoothness of the changes in FA values, length of tracts, and the extent of expected group differences, a useful rule of thumb might be to resample approximately once per voxel. Therefore, for 2.5 mm isotropic data, and a tract that averages 100 mm in length across subjects, we would resample at 40 locations along its length.

The major assumption in doing this, however, is that the ends of all the streamlines – and the vertices that make up the collection at each cross-section – are analogous to each other and therefore appropriate to lump together during statistical comparisons. Radial variability goes against this assumption, but, as previously discussed, this is relatively minor compared to the longitudinal variability. This assumption may also be less valid towards the ends of a tract group, where the individual streamlines can stray to different terminal areas, for example in the corticospinal tract, where most streamlines terminate in the lowest slice of the brainstem, but the most inferior cross-section will also include the final vertices of a few streamlines that strayed posteriorly into the cerebellum. However, with many streamlines making up each tract (approximately 100 per tract on average in our sample analysis), a few spurious fibers are unlikely to have much impact on individual subjects' along-tract estimates. Nevertheless, these issues should always be considered during statistical analysis, as varying noise levels between groups can alter their power to detect effects, and thus potentially lead to spurious conclusions. So far we have described an approach with higher *sensitivity*, which might be a useful way to explore near-cortical white matter at the expense of higher variability in the observations. Alternatively,

for an approach with higher *specificity*, one could apply the same workflow we have described, but limit the analysis to more specific portions of the tract. This could be performed by extracting a subsection of the tract during delineation, subjecting the streamlines to additional constraints (length filters, uncertainty measures, etc.), or by restricting the along-tract analysis to only the high confidence central portions of tracts.

In spite of these shortcomings, this assumption of cross-sectional uniformity is more appropriate than assuming that *all* the vertices in the *entire* tract are analogous and comparable — as is the case with the tract-averaged approach. It is also important to note that this is more of a registration issue, and would still need to be considered carefully with more complex within-tract methods that might look at shape metrics or the tensor fields themselves. Indeed, even though tractography can circumvent some of the issues of traditional voxelwise registration – since the tract groups are individualized dissections based on the white matter anatomy of each subject – on some level you are still going to have to make the assumption that what you are trying to compare between individuals *should* actually be compared.

Other along-tract approaches

Instead of pinning the streamlines together at either end, another resampling strategy is to pin them down at some place in the middle of the tract. This is the approach taken by Corouge et al. (2006), where a manually identified slice plane is used to prescribe corresponding origins across streamlines somewhere in the middle of the tract where there is presumably higher confidence. While this might be advantageous for some tract geometries (In Fig. 11, panel D appears more appropriate than panel B because it avoids the partial-volume-like averaging of possibly different vertex populations towards the ends of the streamlines), it might be less appropriate for others (In Fig. 11, panel A appears more appropriate than panel C because it maintains the correspondence of vertices according to the angular geometry of the tract). This also becomes challenging for between-subject statistical analysis, given that between-subject scaling must be implemented, and there are issues of how to deal with differing numbers of vertices between streamlines and subjects (Corouge et al., 2006). To address the issue of differing tract scale among individuals when parameterizing by arc length, others have used deformable registration to bring individual subjects' tract shapes into alignment with atlas-based templates, removing shape variability while still allowing a type of along-tract parameterization (Goodlett et al., 2008). However, since the sample locations along the streamlines are still given by a continuous arc length variable (albeit now a *standardized* arc length) there is a remaining correspondence issue since there is not a one-to-one mapping of vertices across streamlines and subjects. One progressive approach to circumvent this challenge has been to model the underlying continuous biological variation in FA *directly* though a higher dimensional framework for statistical inference called functional data analysis (Goodlett et al., 2008, 2009; Zhu et al., 2010, 2011). A simpler general approach, which has also been successful, only considers a more proximal set of vertices that have arc-length positional correspondence across all streamlines. This correspondence has variably been determined by 1) moving outwards with constant spacing from some central origin as shown in Figs. 11C,D, or 2) updating the group of corresponding vertices at each point along a mean (or otherwise prototypical) fiber geometry by considering the Euclidean distance of fiber points to each point on the prototype (Maddah et al., 2008) or other optimized cost metrics (O'Donnell et al., 2009). While these methods provide specificity, they sacrifice sensitivity to effects in the more distal regions of the longer streamlines that are effectively ignored. To gain access to these more variable (but possibly interesting) regions, an alternative approach is to group all of the continuous arc length positions into a set of discrete bins, which can then be analyzed as a factor-coded variable (Madden et al., 2009a, 2009b; O'Donnell et al., 2009). As a default option, we make available a somewhat hybrid approach in our package: an extra tie-down origin can be implemented in the middle of the tract by 1) determining the mean tract geometry (see Extract scalars and collapse cross-sectionally section), 2) assigning correspondence to the vertex in each streamline that is closest to the midpoint of this mean tract geometry, and 3) resampling the streamlines with a constant number of vertices, and with an equal proportion lying on either side of the tract midpoint. This allows for the prescription of an interior correspondence point, as suggested by Corouge et al. (2006), but does so in a way that is fully automated and avoids the extra resampling step associated with re-binning an already-interpolated arc length paramterization. Further, this approach utilizes information provided by a kind of prototypical fiber geometry (Maddah et al., 2008; O'Donnell et al., 2009), but does so in a way that still retains sensitivity to the most distal parts of the tract groups. Fig. 9A shows the evolution of this correspondence scheme as an example corticospinal tract fiber group progresses through the along-tract processing workflow. These "correspondence plots", in the efficient standardized style of O'Donnell et al. (2009), use color to show precisely which vertices will be grouped together and compared in the final analysis. This can be helpful to users of the along-tract workflow by explicitly demonstrating the effects that various processing choices have on the point alignment and correspondence scheme of their own tract groups.

To maintain balance between accessibility and modeling complexity, we employed the "constant number of vertices" resampling strategy. However, because this workflow is open source and modular, one has this flexibility to easily incorporate other variations on the individual processing components into these broader themes. For instance, consider the intuitive techniques of using functional Brodmann area masks (Oh et al., 2009) or the "bend" in the arcuate fasciculus (Yeatman et al., 2012) to prescribe additional points of correspondence across streamlines. These approaches could be easily utilized through the current framework, while retaining the advantages of fiber-tracking software interoperability, automation across multiple subjects and tract files, and rich analysis and visualization facilities. For a more *sensitive* style, the user can input the full tract groups, while for a highly *specific* style, the user can prescribe an additional interior origin, and choose to clip their tract groups using the "cut" strategy. To demonstrate the extensibility of our workflow, and further validate our hybrid resampling scheme

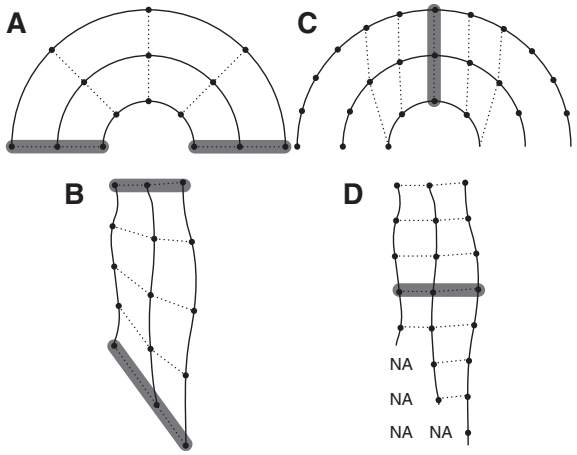


Fig. 11. Comparison of resampling strategies: (A,B) Streamlines resampled with a constant number of vertices, but variable spacing, and pinned together at either end (tract origins designated by the gray groupings of vertices). (C,D) Streamlines resampled with variable numbers of vertices, but constant spacing, and pinned together at the midpoint of the mean tract geometry.

against existing methods, we also implemented the "Distance Map" (Maddah et al., 2008) and "Optimal Point" (O'Donnell et al., 2009) correspondence schemes from the literature. By integrating these alternate algorithms in the context of the generalized along-tract workflow described in this report, the user is given immediate access the tools that are already available for existing methods. For example, we showed how this could be used to generate faceted figures of correspondence plots (Fig. 9), as well as fully-automated, richly detailed, along-tract plots of the extracted data (Fig. 10). This is not meant as an exhaustive exploration of these different options, as there will be some variation in the data extracted by these methods as their different parameters are adjusted. Rather, this comparison shows that all methods are generally successful at extracting along-tract data from tractography fiber groups, and all benefit from being implemented in a broader along-tract workflow.

Example analyses

It is clear from applying these along-tract methods to example data that they offer promising advantages for many tractography applications. Rather than treating each tract as homogenous, these methods reveal significant along-tract variations in FA for all of the major white matter tracts studied, as would be expected based on previous voxel-based studies of the white matter (Schmithorst and Yuan, 2010; Wozniak and Lim, 2006), and other tract-based approaches that have examined FA within tracts (Concha et al., 2010; Davis et al., 2009; Sullivan and Pfefferbaum, 2006; Xue et al., 1999; Yushkevich et al., 2008). The largest benefits of shifting to an along-tract analysis are seen in the subset of tracts that are long and fairly rope-like (Fig. 5), including the corticospinal tract, the inferior fronto-occipital fasciculus, and the occipital projections of the corpus callosum (i.e., the *forceps major*). Here, along-tract modeling gives a large improvement in detail and a prominent decrease in the residual unexplained variance components (e.g., in the corticospinal tract, the tract-averaged standard deviation of 0.16 decreases to less than 0.08 at many of the along-tract cross-sections). This also suggests that along-tract methods may be useful in graph-theory-type structural brain network studies, as these analyses typically employ the type of small, sometimes distant, tracking ROIs that may benefit most from this approach (Bullmore and Sporns, 2009; Hagmann et al., 2010). Further, in these analyses we commonly scale the number of streamlines between two ROIs by some measure of their quality – like average FA or inverse diffusivity – in order to investigate interesting network properties like efficiency (Rubinov and Sporns, 2010). Along-tract methods may be able to contribute to such studies by allowing for their extension to investigate focal along-tract hypotheses that are born out of an initial network analysis.

Another benefit of along-tract methods is that they can be used to address some types of partial volume effects by implementing quality control measures to check the along-tract streamline distribution for outliers or multiple streamline populations. For example, the symmetric parasagittal dips in FA in the *forceps major* likely result from partial volume averaging with cerebrospinal fluid in the adjacent lateral ventricles (Jones et al., 2005). Similarly, along-tract techniques might also be used to highlight crossing fibers areas that are not resolved by the single tensor model of diffusion.

The example between-group analysis demonstrates the richness with which these methods allow one to interrogate their data (Fig. 6). The finding of decreased FA in the inferior longitudinal fasciculus of the FASD group is particularly useful to demonstrate the advantages of modeling diffusion indices along tracts, as the effect is rather localized to only a portion of the entire tract. Although this effect region might generate a significant tract-averaged finding with enough subjects, it could easily be missed if the study were not sufficiently powered, or if there were intermingled effects in the opposite direction. Thus, the along-tract analysis increases power to detect these more focal effects, and also provides better ability to *localize* effects. For example, in regards to fetal alcohol exposure, an along-tract finding in this part of the ILF is consistent with a previous tract-averaged finding in the same tract (Lebel et al., 2008a), but is also consistent with previous voxelwise findings of decreased FA in FASD subjects in the posterior temporal lobe (Sowell et al., 2008). This simple meta-analytical comparison becomes more meaningful with the addition of within-tract detail, as previously it might have been ambiguous whether similar effects reported in a tractography study and a voxelwise study were actually localized to similar regions.

Other options for improving within-tract tractography detail

Segmentation of fiber tracts

Segmentation is perhaps the simplest method to achieve greater within-tract detail, but has also had some of the broadest impact because of its easy implementation and long history of effectiveness in T1-weighted anatomical analyses. The prototypical example of this approach is an analysis of the midline corpus callosum, where it is segmented into different sub-regions that are analyzed individually. These segmentations may be generated, for example, by straightforward coronal slice planes (Wozniak et al., 2009), but have also been created by a variety of other strategies. Even before the widespread adoption of diffusion imaging, this general approach allowed for regional summaries to be generated of voxelwise and morphological analyses (e.g., (Riley et al., 1995; Thompson et al., 2000)). More recently, segmented tractography analyses have been attained in a similar manner by employing sets of masks to extract tract groups that originate from different anatomical portions of the corpus callosum (Schulte et al., 2010), or alternatively, project to different functional regions of the cortex (Huang et al., 2005; Lebel et al., 2010; Whitford et al., 2010). In an interesting variation on this theme, a prescribed structural or functional segmentation of the streamlines at one tract terminus can even be used to drive a connectivity-based segmentation of the structures that lie at the opposite tract terminus (Behrens et al., 2003). Compared to the types of methods described in this report that focus on the *along*-streamline variance component, this approach is typically implemented as a way to model the *between*-streamline variance component without having to resort to a more comprehensive within-tract framework (Yushkevich et al., 2008; Zhang et al., 2010).

Tract-based masks for voxelwise analysis

Instead of using different voxelwise landmarks to restrict the tractography, as described in the previous section, the process can be performed in a somewhat inverse manner, with the tract groups used to mask a voxelwise analysis, either voxel-by-voxel within the tract group mask, or collapsed along a single anatomical axis. This latter approach, in particular, is a close relative of the along-tract methods described here, and reveals some of the along-tract variation, particularly for relatively linear tracts that parallel one of the anatomical axes (e.g., the corticospinal tract) (Pfefferbaum et al., 2005; Sullivan and Pfefferbaum, 2006; Wakana et al., 2007; Xue et al., 1999). However, for tracts with more complex geometries that don't align linearly along one of these axes (e.g., the *forceps major*/*minor*, arcuate fasciculus, etc.), along-tract parameterization like we have described is needed to uncover these variations.

Shape analysis

Building on the long history of surface-based anatomical analyses of T1-weighted data (Fischl and Dale, 2000; Luders et al., 2004; Sowell et al., 2003), the *shapes* of the white matter tracts can provide important complementary information to the voxelwise intensity values that lie within the tract groups. Shape information can be used to drive a registration algorithm, bringing voxelwise (Eckstein et al., 2009) or deterministic/probabilistic-based tract groups into alignment with a common template model (Clayden et al., 2007),

and allowing voxelwise analysis of intensity values throughout their interior, or projection of these values onto the shape-based tract model and subsequent surface-based analysis. Compared to the simpler spline-based along-tract strategies, which assume cross-sectional symmetry along the tracts, these more generic within-tract methods capture variability along more dimensions within a tract (Yushkevich et al., 2008; Zhang et al., 2010). While this might not be an important distinction for the long rope-like tract groups that are typical of tract dissections towards distinct functional areas (e.g., the corticospinal tract projections to the *primary motor cortex*), it could be particularly valuable as a way to investigate variability across related bundles (e.g., the *entire* fanning geometry of the corticospinal tract, or the *entire* set of fibers that pass anywhere through the corpus callosum). An alternative approach to utilize shape information is to directly analyze the shape properties, for example, the vertex-wise deformation needed to bring each tract group shape into register (Qiu et al., 2010), the shape "context" contributed by analyzing where a set of streamlines travel beyond a voxel of interest (Adluru et al., 2009), or streamline curvature and torsion (Batchelor et al., 2006). In the context of the present report, these types of metrics can be mapped to the vertex-wise tract locations to provide complementary information to FA. For instance, the relative size (i.e. width) of the tract along its path could be directly calculated from the spatial spread of the vertices at each cross-section. While lacking a framework for inference, early work by Jones et al. (2005) to map alternative diffusion tensor indices and fiber orientation uncertainty *along* tracts demonstrated the rich potential of pursuing such integration.

Directionality-aware registration

A final class of advanced methods incorporate some sort of directionality information to drive the registration of the voxel-wise diffusion imaging data. This has been performed by developing methods to process the diffusion tensors (Alexander et al., 2001; Arsigny et al., 2006; Corouge et al., 2006; Yeo et al., 2009; Zhang et al., 2006) or the higher order orientation distribution function (ODF) representations of high angular resolution diffusion imaging (HARDI) reconstruction schemes (Chiang et al., 2008), but has also been achieved by either splitting the tensor data into multiple channels (Park et al., 2003), or summarizing the tensor data in single metrics that can be incorporated into the cost functions of existing registration algorithms in a simpler manner (Yap et al., 2009).

Relation to other diffusion models

While this report has focused on example data from a clinical-type DTI sequence (30 directions, 2.5 mm isotropic voxels) analyzed with the FACT algorithm, the general principles can extend to a variety of other diffusion models and tractography approaches that can similarly result in an analogous set of polyline-based tract groups, and their utility is likely to grow with advancements that continue to increase the effective resolution of diffusion imaging data, as this will only further highlight the within-tract heterogeneity. Similarly, these analysis tools are equally relevant to the study of additional diffusion imaging indices like mean diffusivity or axial and radial diffusion components. Other types of data could also be integrated — for example, maps could be generated showing where frontal lobe measures of executive functioning correlate with diffusion imaging indices along tracts leading from the frontal lobe. Perhaps most interesting, other types of imaging data could even be mapped to these same anatomical locations in a multi-modal approach that could examine correlations between along-tract estimates of diffusion imaging indices, and things like local shape attributes, adjacent cortical thickness, sulcal depth, or even fMRI activation indices.

Conclusion

It is clear from inspection of deterministic tractography dissections that there are prominent variations in scalar diffusion imaging metrics (like FA) within the major white matter tracts in the human brain. However, the majority of diffusion tractography analyses still rely on a whole-tract-averaged approach for analyzing differences in these scalar metrics. Moreover, despite excellent work on individual topics related to along-tract processing, and promising advancements towards even more exotic forms of statistical inference in this arena, this present work provides the first extensible end-to-end along-tract workflow for performing and robustly visualizing a standard multiple regression GLM analysis, and an accompanying primer focused on practical considerations for applications-oriented scientists and clinicians. By assuming tracts are relatively tube-like structures with cross-sectional uniformity, we have implemented a straightforward spline-based resampling strategy that captures the large portion of within-tract variance that exists *along* the tracts. These tools have been incorporated into an open source end-to-end workflow for along-tract analysis, and are important because they: 1) Easily integrate into existing tractography studies, 2) Reveal a much richer data landscape than what is typically utilized by traditional tractography methods, 3) Directly extend to enable flexible between-group statistical analyses, and 4) Offer the opportunity to enhance the connectivity analyses of a wide range of neuroscience applications.

Acknowledgments

This work was performed in conjunction with the Collaborative Initiative on Fetal Alcohol Spectrum Disorders (CIFASD), which is funded by grants from the National Institute on Alcohol and Alcohol Abuse (NIAAA). Additional information about CIFASD can be found at www.cifasd.org. This work was also supported by the following organizations: National Institute of Child Health and Human Development (NICHD; R01 HD053893), National Center for Research Resources (NCRR; U54 RR021813, P41 RR013642), National Institute of General Medical Sciences (NIGMS; T32 GM008042), National Institute on Drug Abuse (NIDA; R01 DA017830, R90 DA023422), National Institute on Alcohol Abuse and Alcoholism (NIAAA; U01 AA017122, F30 AA020431), National Institute of Mental Health (NIMH; R01 MH087563), and March of Dimes (6-FY2008-50). The authors report no conflicts of interest. Traffic map courtesy of SigAlert, a Westwood One company (http://www.sigalert.com).

References

Adluru, N., Hinrichs, C., Chung, M.K., Lee, J.-E., Singh, V., Bigler, E.D., Lange, N., Lainhart, J.E., Alexander, A.L., 2009. Classification in DTI using shapes of white matter tracts. Conf. Proc. IEEE Eng. Med. Biol. Soc. 2009, pp. 2719–2722.

Alexander, D.C., Pierpaoli, C., Basser, P.J., Gee, J.C., 2001. Spatial transformations of diffusion tensor magnetic resonance images. IEEE Trans. Med. Imaging 20, 1131–1139.

Aron, A.R., Behrens, T.E., Smith, S., Frank, M.J., Poldrack, R.A., 2007. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J. Neurosci. 27, 3743–3752.

Arsigny, V., Fillard, P., Pennec, X., Ayache, N., 2006. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn. Reson. Med. 56, 411–421.

Asato, M.R., Terwilliger, R., Woo, J., Luna, B., 2010. White matter development in adolescence: a DTI study. Cereb. Cortex 20, 2122–2131.

Ashtari, M., Cottone, J., Ardekani, B.A., Cervellione, K., Szeszko, P.R., Wu, J., Chen, S., Kumra, S., 2007. Disruption of white matter integrity in the inferior longitudinal fasciculus in adolescents with schizophrenia as revealed by fiber tractography. Arch. Gen. Psychiatry 64, 1270–1280.

Barnea-Goraly, N., Lotspeich, L.J., Reiss, A.L., 2010. Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract-based spatial statistics. Arch. Gen. Psychiatry 67, 1052–1060.

Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A., 2000. In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. 44, 625–632.

Batchelor, P.G., Calamante, F., Tournier, J.-D., Atkinson, D., Hill, D.L.G., Connelly, A., 2006. Quantification of the shape of fiber tracts. Magn. Reson. Med. 55, 894–903.

Bava, S., Thayer, R., Jacobus, J., Ward, M., Jernigan, T.L., Tapert, S.F., 2010. Longitudinal characterization of white matter maturation during adolescence. Brain Res. 1327, 38–46.
Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A.M., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat. Neurosci. 6, 750–757.
Bengtsson, S.L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F., 2005. Extensive piano practicing has regionally specific effects on white matter development. Nat. Neurosci. 8, 1148–1150.
Boorman, E.D., O'Shea, J., Sebastian, C., Rushworth, M.F.S., Johansen-Berg, H., 2007. Individual differences in white-matter microstructure reflect variation in functional connectivity during choice. Curr. Biol. 17, 1426–1431.
Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.
Chiang, M.C., Leow, A.D., Klunder, A.D., Dutton, R.A., Barysheva, M., Rose, S.E., McMahon, K.L., de Zubicaray, G.I., Toga, A.W., Thompson, P.M., 2008. Fluid registration of diffusion tensor images using information theory. IEEE Trans. Med. Imaging 27, 442–456.
Childs, C., Hiltunen, Y., Vidyasagar, R., Kauppinen, R.A., 2007. Determination of regional brain temperature using proton magnetic resonance spectroscopy to assess brain-body temperature differences in healthy human subjects. Magn. Reson. Med. 57, 59–66.
Clayden, J.D., Storkey, A.J., Bastin, M.E., 2007. A probabilistic model-based approach to consistent white matter tract segmentation. IEEE Trans. Med. Imaging 26, 1555–1561.
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. J. Neurosci. 30, 996–1002.
Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E., 1999. Tracking neuronal fiber pathways in the living human brain. Proc. Natl. Acad. Sci. U. S. A. 96, 10422–10427.
Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G., 2006. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. Med. Image Anal. 10, 786–798.
Davis, S.W., Dennis, N.A., Buchler, N.G., White, L.E., Madden, D.J., Cabeza, R., 2009. Assessing the effects of age on long white matter tracts using diffusion tensor tractography. Neuroimage 46, 530–541.
Dougherty, R.F., Ben-Shachar, M., Deutsch, G.K., Hernandez, A., Fox, G.R., Wandell, B.A., 2007. Temporal–callosal pathway diffusivity predicts phonological skills in children. Proc. Natl. Acad. Sci. U. S. A. 104, 8556–8561.
Duncan, J.S., 2010. Imaging in the surgical treatment of epilepsy. Nat. Rev. Neurol. 6, 537–550.
Eckstein, I., Shattuck, D., Stein, J., McMahon, K., de Zubicaray, G., Wright, M., Thompson, P., Toga, A., 2009. Active fibers: matching deformable tract templates to diffusion tensor images. Neuroimage 47, T82–T89.
Eluvathingal, T.J., Hasan, K.M., Kramer, L., Fletcher, J.M., Ewing-Cobbs, L., 2007. Quantitative diffusion tensor tractography of association and projection fibers in normally developing children and adolescents. Cereb. Cortex 17, 2760–2768.
Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. Natl. Acad. Sci. U. S. A. 97, 11050–11055.
Glenn, O.A., Ludeman, N.A., Berman, J.I., Wu, Y.W., Lu, Y., Bartha, A.I., Vigneron, D.B., Chung, S.W., Ferriero, D.M., Barkovich, A.J., Henry, R.G., 2007. Diffusion tensor MR imaging tractography of the pyramidal tracts correlates with clinical motor function in children with congenital hemiparesis. AJNR Am. J. Neuroradiol. 28, 1796–1802.
Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G., 2008. Group statistics of DTI fiber bundles using spatial functions of tensor measures. Med. Image Comput. Comput. Assist. Interv. 11, 1068–1075.
Goodlett, C.B., Fletcher, P.T., Gilmore, J.H., Gerig, G., 2009. Group analysis of DTI fiber tract statistics with application to neurodevelopment. Neuroimage 45, S133–S142.
Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V.J., Meuli, R., Thiran, J.-P., Grant, P.E., 2010. White matter maturation reshapes structural connectivity in the late developing human brain. Proc. Natl. Acad. Sci. U. S. A. 107, 19067–19072.
Hsu, J., Hecke, W.V., Bai, C., Lee, C., Tsai, Y., Chiu, H., Jaw, F., Hsu, C., Leu, J., Chen, W., Leemans, A., 2010. Microstructural white matter changes in normal aging: a diffusion tensor imaging study with higher-order polynomial regression models. Neuroimage 49, 32–43.
Huang, H., Zhang, J., Jiang, H., Wakana, S., Poetscher, L., Miller, M.I., van Zijl, P.C.M., Hillis, A.E., Wytik, R., Mori, S., 2005. DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. Neuroimage 26, 195–205.
Huang, H., Zhang, J., Wakana, S., Zhang, W., Ren, T., Richards, L.J., Yarowsky, P., Donohue, P., Graham, E., van Zijl, P.C.M., Mori, S., 2006. White and gray matter development in human fetal, newborn and pediatric brains. Neuroimage 33, 27–38.
Innocenti, G.M., Clarke, S., Koppel, H., 1983. Transitory macrophages in the white matter of the developing visual cortex. II. Development and relations with axonal pathways. Brain Res. 313, 55–66.
Ishii, K., Sasaki, M., Kitagaki, H., Sakamoto, S., Yamaji, S., Maeda, K., 1996. Regional difference in cerebral blood flow and oxidative metabolism in human cortex. J. Nucl. Med. 37, 1086–1088.
Ito, H., Kanno, I., Fukuda, H., 2005. Human cerebral circulation: positron emission tomography studies. Ann. Nucl. Med. 19, 65–74.
Jones, D.K., Simmons, A., Williams, S.C., Horsfield, M.A., 1999. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. Magn. Reson. Med. 42, 37–41.
Jones, D.K., Travis, A.R., Eden, G., Pierpaoli, C., Basser, P.J., 2005. PASTA: pointwise assessment of streamline tractography attributes. Magn. Reson. Med. 53, 1462–1467.
Keller, T.A., Just, M.A., 2009. Altering cortical connectivity: remediation-induced changes in the white matter of poor readers. Neuron 64, 624–631.
Kumar, A., Sundaram, S.K., Sivaswamy, L., Behen, M.E., Makki, M.I., Ager, J., Janisse, J., Chugani, H.T., Chugani, D.C., 2010. Alterations in frontal lobe tracts and corpus callosum in young children with autism spectrum disorder. Cereb. Cortex 20, 2103–2113.
Kunimatsu, A., Aoki, S., Masutani, Y., Abe, O., Mori, H., Ohtomo, K., 2003. Three-dimensional white matter tractography by diffusion tensor imaging in ischaemic stroke involving the corticospinal tract. Neuroradiology 45, 532–535.
Le Bihan, D., 2003. Looking into the functional architecture of the brain with diffusion MRI. Nat. Rev. Neurosci. 4, 469–480.
Lebel, C., Beaulieu, C., 2009. Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. Hum. Brain Mapp. 30, 3563–3573.
Lebel, C., Rasmussen, C., Wyper, K., Walker, L., Andrew, G., Yager, J., Beaulieu, C., 2008a. Brain diffusion abnormalities in children with fetal alcohol spectrum disorder. Alcohol. Clin. Exp. Res. 32, 1732–1740.
Lebel, C., Walker, L., Leemans, A., Phillips, L., Beaulieu, C., 2008b. Microstructural maturation of the human brain from childhood to adulthood. Neuroimage 40, 1044–1055.
Lebel, C., Caverhill-Godkewitsch, S., Beaulieu, C., 2010. Age-related regional variations of the corpus callosum identified by diffusion tensor tractography. Neuroimage 52, 20–31.
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. Cereb. Cortex 16, 553–560.
Luck, D., Buchy, L., Czechowska, Y., Bodnar, M., Pike, G.B., Campbell, J.S.W., Achim, A., Malla, A., Joober, R., Lepage, M., 2011. Fronto-temporal disconnectivity and clinical short-term outcome in first episode psychosis: A DTI-tractography study. J. Psychiatr. Res. 45, 369–377.
Luders, E., Narr, K.L., Thompson, P.M., Rex, D.E., Jancke, L., Steinmetz, H., Toga, A.W., 2004. Gender differences in cortical complexity. Nat. Neurosci. 7, 799–800.
Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M., 2008. A unified framework for clustering and quantitative analysis of white matter fiber tracts. Med. Image Anal. 12, 191–202.
Madden, D.J., Bennett, I.J., Song, A.W., 2009a. Cerebral white matter integrity and cognitive aging: contributions from diffusion tensor imaging. Neuropsychol. Rev. 19, 415–435.
Madden, D.J., Spaniol, J., Costello, M.C., Bucur, B., White, L.E., Cabeza, R., Davis, S.W., Dennis, N.A., Provenzale, J.M., Huettel, S.A., 2009b. Cerebral white matter integrity mediates adult age differences in cognitive performance. J. Cogn. Neurosci. 21, 289–302.
Maruyama, K., Kamada, K., Shin, M., Itoh, D., Aoki, S., Masutani, Y., Tago, M., Kirino, T., 2005. Integration of three-dimensional corticospinal tractography into treatment planning for gamma knife surgery. J. Neurosurg. 102, 673–677.
McIntosh, A.M., Moorhead, T.W.J., Job, D., Lymer, G.K.S., Muñoz Maniega, S., McKirdy, J., Sussmann, J.E.D., Baig, B.J., Bastin, M.E., Porteous, D., Evans, K.L., Johnstone, E.C., Lawrie, S.M., Hall, J., 2008. The effects of a neuregulin 1 variant on white matter density and integrity. Mol. Psychiatry 13, 1054–1059.
Moghaddam, B., Adams, R.N., 1987. Regional differences in resting extracellular potassium levels of rat brain. Brain Res. 406, 337–340.
Monier, A., Evrard, P., Gressens, P., Verney, C., 2006. Distribution and differentiation of microglia in the human encephalon during the first two trimesters of gestation. J. Comp. Neurol. 499, 565–582.
Monier, A., Adle-Biassette, H., Delezoide, A.-L., Evrard, P., Gressens, P., Verney, C., 2007. Entry and distribution of microglial cells in human embryonic and fetal cerebral cortex. J. Neuropathol. Exp. Neurol. 66, 372–382.
Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45, 265–269.
Nichols, T., Hayasaka, S., 2003. Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat. Methods Med. Res. 12, 419–446.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
O'Donnell, L., Westin, C., Golby, A., 2009. Tract-based morphometry for white matter group analysis. Neuroimage 45, 832–844.
Oh, J.S., Kubicki, M., Rosenberger, G., Bouix, S., Levitt, J.J., McCarley, R.W., Westin, C.-F., Shenton, M.E., 2009. Thalamo-frontal white matter alterations in chronic schizophrenia: a quantitative diffusion tractography study. Hum. Brain Mapp. 30, 3812–3825.
Pankonin, M.S., Sohi, J., Kamholz, J., Loeb, J.A., 2009. Differential distribution of neuregulin in human brain and spinal fluid. Brain Res. 1258, 1–11.
Park, H.-J., Kubicki, M., Shenton, M.E., Guimond, A., McCarley, R.W., Maier, S.E., Kikinis, R., Jolesz, F.A., Westin, C.-F., 2003. Spatial normalization of diffusion tensor MRI using multiple channels. Neuroimage 20, 1995–2009.
Penke, L., Muñoz Maniega, S., Murray, C., Gow, A.J., Hernández, M.C.V., Clayden, J.D., Starr, J.M., Wardlaw, J.M., Bastin, M.E., Deary, I.J., 2010. A general factor of brain white matter integrity predicts information processing speed in healthy older people. J. Neurosci. 30, 7569–7574.
Perry, T.L., Berry, K., Hansen, S., Diamond, S., Mok, C., 1971. Regional distribution of amino acids in human brain obtained at autopsy. J. Neurochem. 18, 513–519.
Pfefferbaum, A., Adalsteinsson, E., Sullivan, E.V., 2005. Frontal circuitry degradation marks healthy adult aging: Evidence from diffusion tensor imaging. Neuroimage 26, 891–899.
Pinheiro, J., Bates, D., 2000. Mixed-Effects Models in S and S-Plus. Springer, New York.
Prabhu, S.S., Gasco, J., Tummala, S., Weinberg, J.S., Rao, G., 2011. Intraoperative magnetic resonance imaging-guided tractography with integrated monopolar subcortical functional mapping for resection of brain tumors. J. Neurosurg. 114, 719–726.

Qiu, A., Oishi, K., Miller, M.I., Lyketsos, C.G., Mori, S., Albert, M., 2010. Surface-based analysis on shape and fractional anisotropy of white matter tracts in Alzheimer's disease. PLoS One 5, e9811.
Ramnani, N., Behrens, T.E.J., Penny, W., Matthews, P.M., 2004. New approaches for exploring anatomical and functional connectivity in the human brain. Biol. Psychiatry 56, 613–619.
Riley, E.P., Mattson, S.N., Sowell, E.R., Jernigan, T.L., Sobel, D.F., Jones, K.L., 1995. Abnormalities of the corpus callosum in children prenatally exposed to alcohol. Alcohol. Clin. Exp. Res. 19, 1198–1202.
Rubinov, M., Sporns, O., 2010. Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52, 1059–1069.
Sala S., Agosta F., Pagani E., Copetti M., Comi G., and Filippi M., in press. Microstructural changes and atrophy in brain white matter tracts with aging. Neurobiol. Aging.
Schmithorst, V., Yuan, W., 2010. White matter development during adolescence as shown by diffusion MRI. Brain Cogn. 72, 16–25.
Schulte, T., Müller-Oehring, E.M., Rohlfing, T., Pfefferbaum, A., Sullivan, E.V., 2010. White matter fiber degradation attenuates hemispheric asymmetry when integrating visuomotor information. J. Neurosci. 30, 12168–12178.
Shaffer, J.P., 1995. Multiple hypothesis testing. Annu. Rev. Psychol. 46, 561–584.
Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505.
Smith, S.M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T.E., Miller, K.L., Robson, M.D., Jones, D.K., Klein, J.C., Bartsch, A.J., Behrens, T.E.J., 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. Nat. Protoc. 2, 499–503.
Sowell, E.R., Peterson, B.S., Thompson, P.M., Welcome, S.E., Henkenius, A.L., Toga, A.W., 2003. Mapping cortical change across the human life span. Nat. Neurosci. 6, 309–315.
Sowell, E.R., Johnson, A., Kan, E., Lu, L.H., Van Horn, J.D., Toga, A.W., O'Connor, M.J., Bookheimer, S.Y., 2008. Mapping white matter integrity and neurobehavioral correlates in children with fetal alcohol spectrum disorders. J. Neurosci. 28, 1313–1319.
Stamford, J.A., Kruk, Z.L., Millar, J., 1984. Regional differences in extracellular ascorbic acid levels in the rat brain determined by high speed cyclic voltammetry. Brain Res. 299, 289–295.
Stikov, N., Perry, L.M., Mezer, A., Rykhlevskaia, E., Wandell, B.A., Pauly, J.M., Dougherty, R.F., 2011. Bound pool fractions complement diffusion measures to describe white matter micro and macrostructure. Neuroimage 54, 1112–1121.
Sullivan, E.V., Pfefferbaum, A., 2006. Diffusion tensor imaging and aging. Neurosci. Biobehav. Rev. 30, 749–761.
Tabelow, K., Clayden, J.D., de Micheaux, P.L., Polzehl, J., Schmid, V.J., Whitcher, B., 2011. Image analysis and statistical inference in neuroimaging with R. Neuroimage 55, 1686–1693.
Thompson, P.M., Giedd, J.N., Woods, R.P., MacDonald, D., Evans, A.C., Toga, A.W., 2000. Growth patterns in the developing brain detected by using continuum mechanical tensor maps. Nature 404, 190–193.
Tsang, J.M., Dougherty, R.F., Deutsch, G.K., Wandell, B.A., Ben-Shachar, M., 2009. Frontoparietal white matter diffusion properties predict mental arithmetic skills in children. Proc. Natl. Acad. Sci. U. S. A. 106, 22546–22551.
Tufte, E.R., 2001. The Visual Display of Quantitative Information, 2nd ed. Graphics Press, Cheshire, CT.
Vasung, L., Huang, H., Jovanov-Milošević, N., Pletikos, M., Mori, S., Kostović, I., 2010. Development of axonal pathways in the human fetal fronto-limbic brain: histochemical characterization and diffusion tensor imaging. J. Anat. 217, 400–417.
Verhoeven, J.S., Sage, C.A., Leemans, A., Hecke, W.V., Callaert, D., Peeters, R., Cock, P.D., Lagae, L., Sunaert, S., 2010. Construction of a stereotaxic DTI atlas with full diffusion tensor information for studying white matter maturation from childhood to adolescence using tractography-based segmentations. Hum. Brain Mapp. 31, 470–486.
Voineskos, A.N., Rajji, T.K., Lobaugh, N.J., Miranda, D., Shenton, M.E., Kennedy, J.L., Pollock, B.G., Mulsant, B.H., 2012. Age-related decline in white matter tract integrity and cognitive performance: A DTI tractography and structural equation modeling study. Neurobiol. Aging 33, 21–34.
Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., Blitz, A., van Zijl, P., Mori, S., 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36, 630–644.
Warrington, N.M., Woerner, B.M., Daginakatte, G.C., Dasgupta, B., Perry, A., Gutmann, D.H., Rubin, J.B., 2007. Spatiotemporal differences in CXCL12 expression and cyclic AMP underlie the unique pattern of optic glioma growth in neurofibromatosis type 1. Cancer Res. 67, 8588–8595.
Whitford, T.J., Kubicki, M., Schneiderman, J.S., O'Donnell, L.J., King, R., Alvarado, J.L., Khan, U., Markant, D., Nestor, P.G., Niznikiewicz, M., McCarley, R.W., Westin, C.-F., Shenton, M.E., 2010. Corpus callosum abnormalities and their association with psychotic symptoms in patients with schizophrenia. Biol. Psychiatry 68, 70–77.
Wickham, H., 2009. ggplot2 — Elegant Graphics for Data Analysis. Springer, New York.
Wilkinson, L., 2005. The Grammar of Graphics, 2nd ed. Springer, New York.
Wozniak, J.R., Lim, K.O., 2006. Advances in white matter imaging: a review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. Neurosci. Biobehav. Rev. 30, 762–774.
Wozniak, J.R., Muetzel, R.L., Mueller, B.A., McGee, C.L., Freerks, M.A., Ward, E.E., Nelson, M.L., Chang, P.-N., Lim, K.O., 2009. Microstructural corpus callosum anomalies in children with prenatal alcohol exposure: an extension of previous diffusion tensor imaging findings. Alcohol. Clin. Exp. Res. 33, 1825–1835.
Xue, R., van Zijl, P.C., Crain, B.J., Solaiyappan, M., Mori, S., 1999. In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging. Magn. Reson. Med. 42, 1123–1127.
Yap, P.-T., Wu, G., Zhu, H., Lin, W., Shen, D., 2009. TIMER: tensor image morphing for elastic registration. Neuroimage 47, 549–563.
Yeatman, J.D., Dougherty, R.F., Rykhlevskaia, E., Sherbondy, A.J., Deutsch, G.K., Wandell, B.A., Ben-Shachar, M., 2011. Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. J. Cogn. Neurosci 23, 3304–3317.
Yeh, T.-H., Lee, D.Y., Gianino, S.M., Gutmann, D.H., 2009. Microarray analyses reveal regional astrocyte heterogeneity with implications for neurofibromatosis type 1 (NF1)-regulated glial proliferation. Glia 57, 1239–1249.
Yeo, B.T.T., Vercauteren, T., Fillard, P., Peyrat, J.-M., Pennec, X., Golland, P., Ayache, N., Clatz, O., 2009. DT-REFinD: diffusion tensor registration with exact finite-strain differential. IEEE Trans. Med. Imaging 28, 1914–1928.
Young, R.J., Brennan, N., Fraser, J.F., Brennan, C., 2010. Advanced imaging in brain tumor surgery. Neuroimaging Clin. N. Am. 20, 311–335.
Yushkevich, P.A., Zhang, H., Simon, T.J., Gee, J.C., 2008. Structure-specific statistical mapping of white matter tracts. Neuroimage 41, 448–461.
Zarei, M., Damoiseaux, J., Morgese, C., Beckmann, C., Smith, S., Matthews, P., Scheltens, P., Rombouts, S., Barkhof, F., 2009. Regional white matter integrity differentiates between vascular dementia and Alzheimer disease. Stroke 40, 773–779.
Zhang, H., Yushkevich, P.A., Alexander, D.C., Gee, J.C., 2006. Deformable registration of diffusion tensor MR images with explicit orientation optimization. Med. Image Anal. 10, 764–785.
Zhang, H., Awate, S.P., Das, S.R., Woo, J.H., Melhem, E.R., Gee, J.C., Yushkevich, P.A., 2010. A tract-specific framework for white matter morphometry combining macroscopic and microscopic tract features. Med. Image Anal. 14, 666–673.
Zhu, H., Styner, M., Tang, N., Liu, Z., Lin, W., Gilmore, J., 2010. FRATS: Functional regression analysis of DTI tract statistics. IEEE Trans. Med. Imaging 29, 1039–1049.
Zhu, H., Kong, L., Li, R., Styner, M., Gerig, G., Lin, W., Gilmore, J.H., 2011. FADTTS: functional analysis of diffusion tensor tract statistics. Neuroimage 56, 1412–1425.

CHAPTER 6

Fetal Alcohol Spectrum Disorders (FASDs)

6.1 Demographics

Data were collected between Summer 2008 and Summer 2011 according to the protocol described in § 3.3.4. During this time, 70 participants were enrolled for scanning at UCLA. Of these, 61 completed one or more DTI acquisitions without any gross artifacts (*e.g.* aliasing along the anterior–posterior phase-encoding direction, clipping of the top/bottom of the brain, whole-slice dropout due to ballistic head movements). Additional subjects were excluded if they were pilots scans (1), young adults (age > 18) (2), members of patient groups for other studies (5), or longitudinal repeat visits (2). This left 51 total subjects available for the analysis.

According to the diagnostic criteria described in §2.3.1 and §3.4.2, participants were assigned to either the "FASD" or "Control" groups. The control group included 29 subjects (16 female, age 12.5 ± 2.9 years), and the FASD group included 22 subjects (11 female, age 12.9 ± 2.3 years) (Figure 6.1). Groups did not significantly differ in age or gender. Importantly, 10 FASD subjects also had exposure to methamphetamine.

6.2 Voxelwise analysis

Group differences in FA between the FASD and Control groups were assessed with TBSS. The analysis protocol generally followed the descriptions in § 2.3 and Chapter 4. However, here the FMRIB58 FA atlas brain was used as the nonlinear registration target. Using a linear model with group intercept terms and an across-group age covariate, we observed the typical developmental

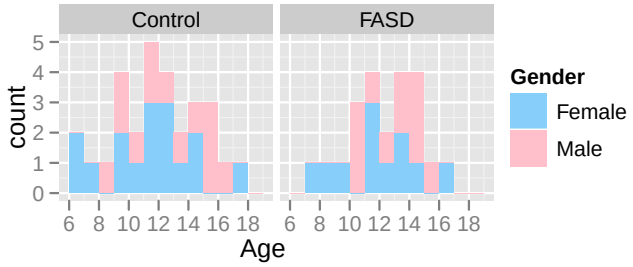


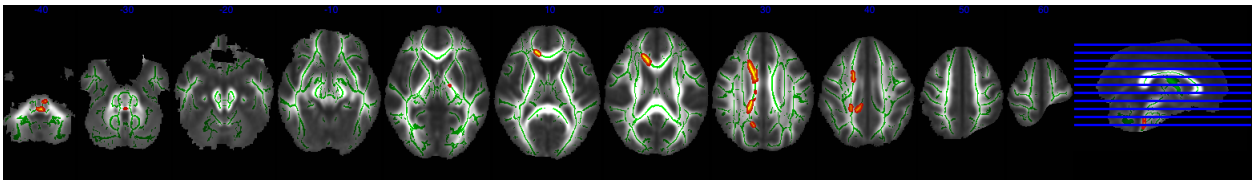
Figure 6.1: Age distributions by gender and group

pattern of increasing FA and decreasing MD, driven by decreasing AD and even larger decreases in RD (Figure 6.2). However, correcting for multiple comparisons, no significant voxels were found for the group contrasts for any of the DTI metrics.

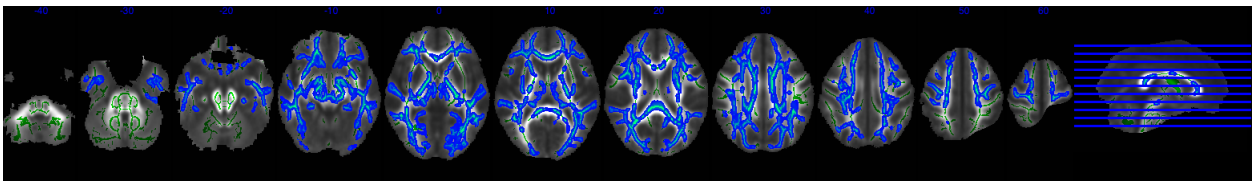
We also failed to find any significant voxels in the group contrasts when this analysis was repeated using the restricted set of only 12 FASD subjects who did not have any comorbid exposure to methamphetamine. The age effects in this follow-up analysis again showed increasing FA, and decreasing MD, AD, and RD. In particular, the regions of positive correlation between FA and age were more widespread in this analysis (Figure 6.3).

6.3 Tractography analysis

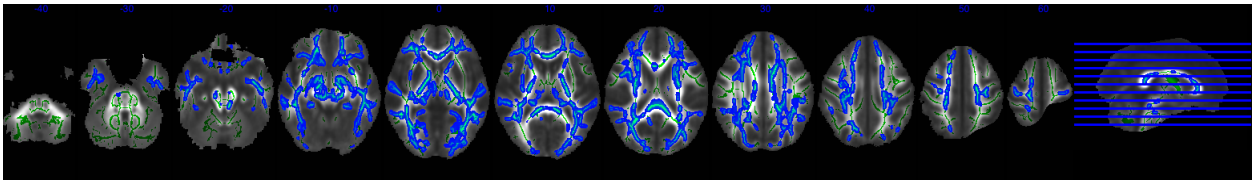
Whole-brain streamline tractography data were generated and analyzed according to the methods described in §3.3.7 and Chapter 5. Briefly, for each subject's preprocessed tensor data, this included fiber-tracking with the standard Fiber Assignment by Continuous Tracking (FACT) algorithm (Mori *et al.*, 1999), manual delineation of 10 major white matter tracts by a trained experimenter (Wakana *et al.*, 2007), tabulation of DTI metrics at many points along these tracts, and between-group statistical modeling (Colby *et al.*, 2012). For an explanation of tract abbreviations, see especially Figure 5 in the Colby *et al.* (2012) paper.



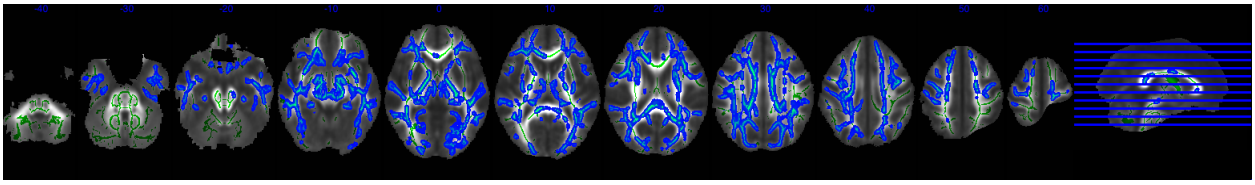
(A) (+) Fractional anisotropy (FA)



(B) (−) Mean diffusivity (MD)



(C) (−) Axial diffusivity (AD)



(D) (−) Radial diffusivity (RD)

Figure 6.2: Voxelwise correlations between DTI metrics and age ($n = 51$, $P < 0.05$).

6.3.1 Across-tract observations

6.3.1.1 Tracking failures

The first step in analyzing deterministic tractography data is to simply tabulate the number of tract groups where tracking was successful (*i.e.* one or more fibers were identified) versus the number of failures (*i.e.* zero fibers were identified) (Table 6.1). The number of failures was highest in the arcuate fasciculus — particularly the right hemisphere. In proportion to their sample sizes, the FASD group also tended to have more tracking failures overall, but this was not statistically significant ($P = 0.15$).

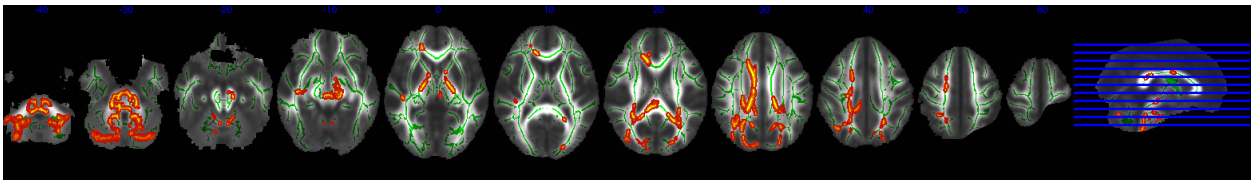


Figure 6.3: Voxelwise correlations between FA and age ($n = 41$, $P < 0.05$).

	Tract	Hemisphere	Control	FASD
1	CGC	L	3.45	4.55
2	CGC	R	6.90	9.09
3	CGH	L	3.45	9.09
4	CGH	R	6.90	9.09
5	CST	L	3.45	4.55
6	CST	R	3.45	4.55
7	ATR	L	3.45	0.00
8	ATR	R	3.45	0.00
9	AF	L	3.45	13.64
10	AF	R	41.38	68.18
11	ILF	L	6.90	4.55
12	ILF	R	3.45	4.55
13	IFO	L	3.45	4.55
14	IFO	R	3.45	4.55
15	UNC	L	3.45	9.09
16	UNC	R	3.45	9.09
17	Fmajor	L	3.45	0.00
18	Fminor	L	3.45	0.00

Table 6.1: Tracking failures (%) by tract, hemisphere, and group

6.3.1.2 Tract length

Tract length naturally varies by tract, so average tract lengths were first calculated and used to determine the number of interpolation points (`nPts`) for extracting along-tract data. The numbers of interpolation points were chosen such that the underlying data get sampled approximately once per voxel. Tract lengths did not significantly vary by hemisphere, so both versions were interpolated with the same number of points to allow for potentially investigating group:hemisphere types of effects. The average tract lengths and numbers of interpolation points are listed in Table 6.2. For the subjects who had greater than zero fibers for a given tract, the FASD group generally had shorter tracts ($P < 0.001$). This effect did not significantly vary by tract ($P = 0.06$).

	Tract	Hemisphere	Mean length (mm)	nPts
1	CGC	L	75.78	29
2	CGC	R	68.29	29
3	CGH	L	35.92	15
4	CGH	R	39.49	15
5	CST	L	102.28	41
6	CST	R	101.75	41
7	ATR	L	65.63	27
8	ATR	R	67.52	27
9	AF	L	84.55	35
10	AF	R	85.37	35
11	ILF	L	83.76	35
12	ILF	R	85.11	35
13	IFO	L	133.86	55
14	IFO	R	135.96	55
15	UNC	L	59.02	25
16	UNC	R	59.56	25
17	Fmajor	L	136.69	55
18	Fminor	L	92.62	37

Table 6.2: Average tract lengths by tract and hemisphere

6.3.1.3 Number of streamlines

Similar to length, the number of streamlines in the tractography dissections varied by tract. When considering all tracts jointly with ANOVA, we did not observe any significant overall hemisphere or hemisphere:tract interactive effects. There was, however, a subthreshold trend towards a main group effect of fewer streamlines on average in the FASD group ($P = 0.10$), as well as tract:group ($P = 0.08$) interactive effects. These distributions are plotted in Figure 6.4 in the right set of panels.

6.3.2 *Along*-tract observations

For each tract and hemisphere, FA was modeled as a function of along-tract position, group (FASD or Control), position:group interactions, and age. A single subject-level random intercept term was also included. Because no significant group effects were observed when the TBSS analysis was corrected for multiple comparisons, here we decided to first focus on the uncorrected maps so at least we could qualitatively explore the trends in the data. Using this model, significant posi-

tion:group interactive effects were identified in the left cigulate gyrus (hippocampal part), bilateral corticospinal tracts, left inferior longitudinal fasciculus, and bilateral inferior frontal occipital fasciculi. These types of effects correspond to focal regions of group differences that are isolated to regions along the tracts. Examining the individual t statistics along the tracts (Figure 6.4) reveals the nature of these effects to be generally lower FA in the FASD group. However, the L ILF is an exception, and has a mixed picture with regions of both higher and lower FA in the FASD group.

The entire along-tract data set is displayed in Figure 6.4. Accompanying t-statistic and P-value plots for the main group contrast are shown in Figure 6.5. For reference, *all* pairwise significant regions are shown in this plot, instead of masking them to only show the panels with a significant position:group F test as in Figure 6.4. P-values and effect sizes are also useful to visualize projected back onto the 3D mean tract geometry of an example subject (Figure 6.6).

For familiarizing oneself with these types of figures, it is useful to first focus on one tract region and then identify it on all of the plots. For example, consider the inferior frontal occipital fasciculus, which is a long tract connecting the frontal lobe (0% position in our along-tract plots) to the occipital lobe (100% position in our along-tract plots) through the external capsule. It has significant position:group interaction effects in both hemispheres, signaling the presence of focal along-tract group effects. Therefore the regions within this tract that have significant pairwise t statistics are tagged with black bars toward the bottom of the panels in Figure 6.4. Inspection of the group mean curves in these same plots tells us that these effects are due to lower FA in the FASD group. These significant regions are also highlighted in green in the analogous panels on Figure 6.5, with the t-statistic plot again showing the directionality of this effect. Finally, moving to the 3D plots (Figure 6.6), these regions show up in blue (*i.e.* negative effects in FASD relative to control subjects) towards the posterior end of IFO.

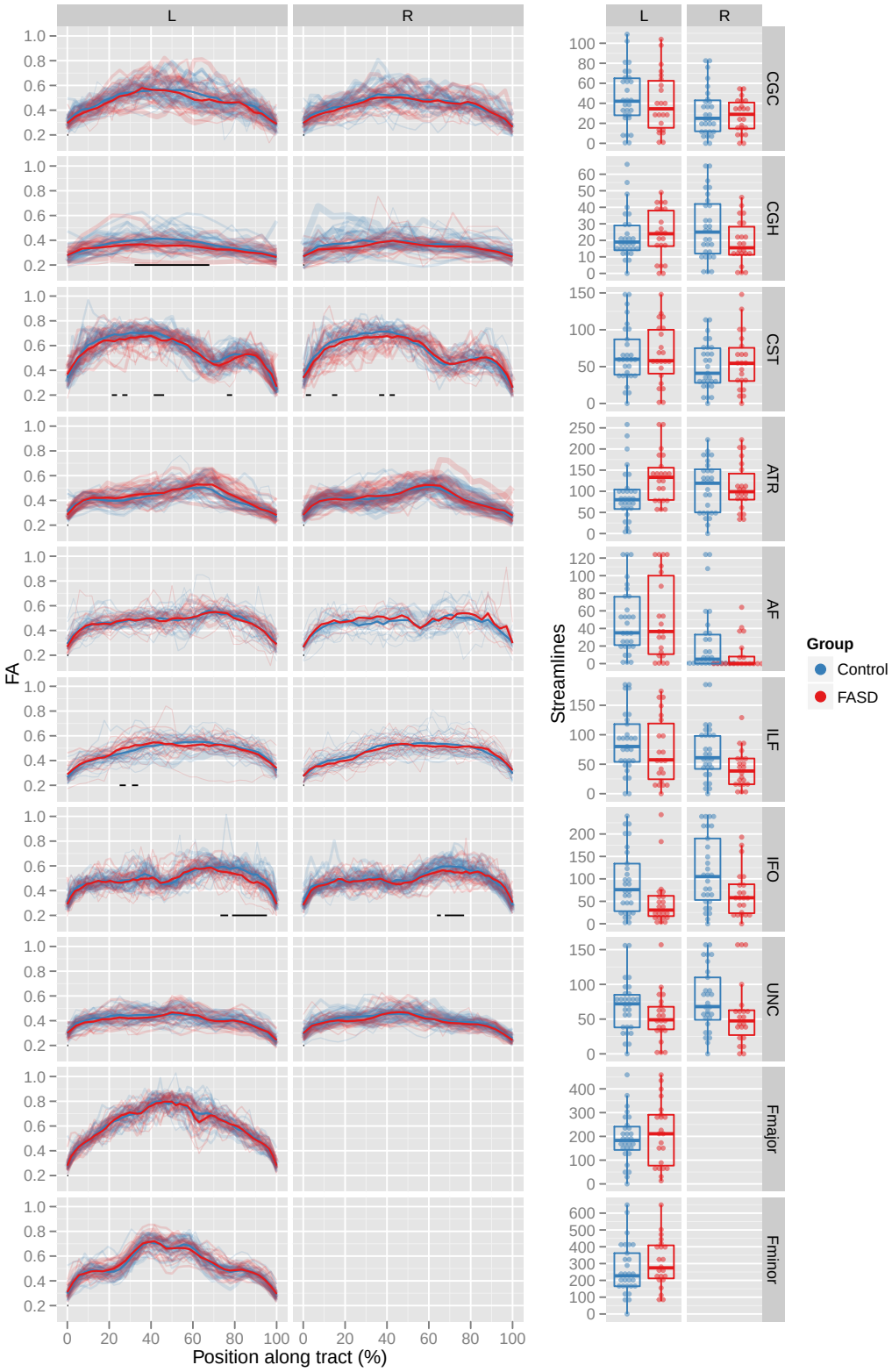
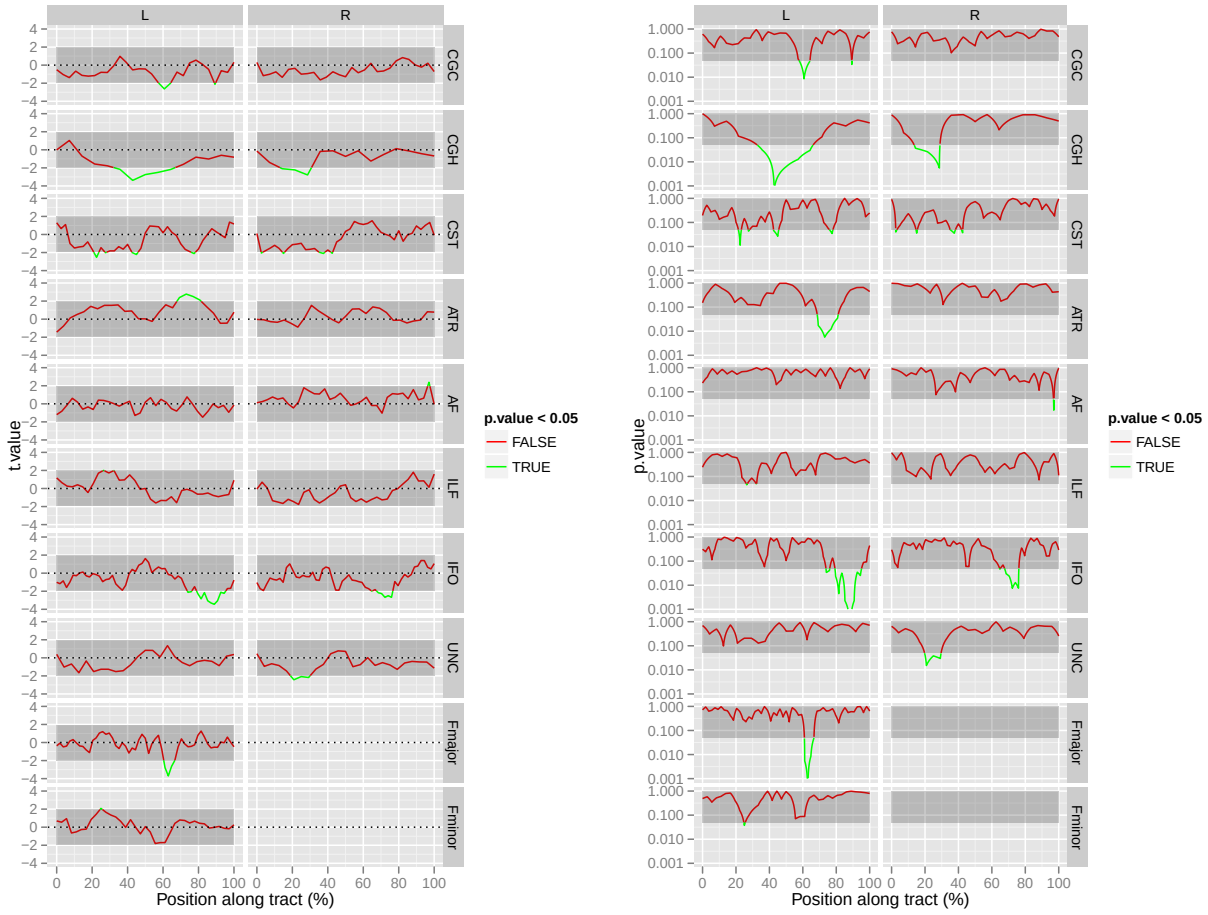
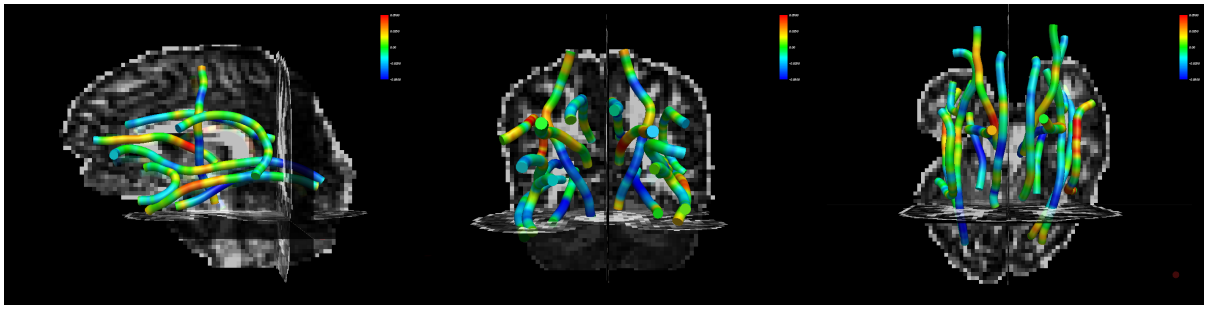

L
R
FA
Position along tract (%)
Streamlines
CGC
CGH
CST
ATR
AF
ILF
IFO
UNC
Fmajor
Fminor
Group
Control
FASD

Figure 6.4: Along-tract variation in FA, by tract, hemisphere, and group. FA is plotted versus position along tract, and panels are facetted into a grid by tract and hemisphere. 0% typically corresponds to the frontal lobe tract terminus (exceptions: CST, inferior terminus; CC, left hemisphere terminus). Color encodes group membership, and line-width encodes the number of streamlines for a given subject. Individual subjects are displayed semi-transparently in the background, and group means are overlaid on top. For panels with a significant position:group interaction, significant regions of group effects are highlighted towards the bottom of each panel. Distributions of the number of streamlines are shown in an accompanying set of panels at right.

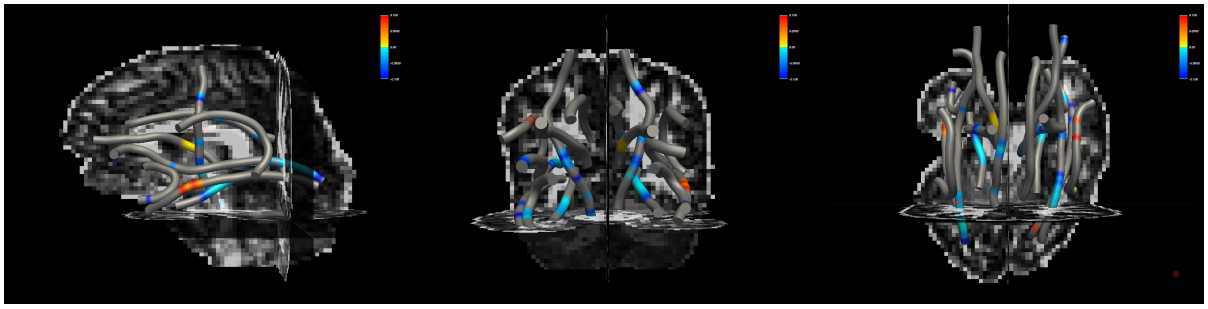


(A) *t*-statistic plot **(B)** *P*-value plot

Figure 6.5: 2D statistical plots for Figure 6.4 (FASD–Control contrast)



(A) Effect size (FA in FASD – Control)



(B) 2-way *P*-value plot

Figure 6.6: 3D statistical plots for Figure 6.4 (FASD–Control contrast)

6.4 Graph-based network analysis

In additional to the voxelwise and tract-based brain mapping approaches, which are serially univariate with respect to different anatomical locations, it is reasonable that analyzing the *multivariate* patterns of structural abnormalities *across* regions, and their relation to cognitive measures, might provide additional and complementary insight into the biological basis of typical brain development and its perturbation in the context of FASDs.

Graph-theory-based network analysis is one such approach that has recently emerged as a popular tool in the neuroimaging community (Bullmore and Sporns, 2009; Dosenbach *et al.*, 2010; Fair *et al.*, 2008). This is in part because the field has a long and successful history of theoretical advancements, as well as applications to the study of topological properties of other physical/communication/information systems (Strogatz, 2001). In the language of graph analysis, the network topology of an individual brain's structural or functional network connectivity can be summarized as a list of nodes (*i.e.* focal brain regions) and their pairwise connection strengths.

The physical/imaging metric prescribed to this connectedness quality is variable, and, for example, could be the number of streamline fibers between nodes in a deterministic tractography analysis (Brown *et al.*, 2011), or the pairwise time series correlation in a functional connectivity analysis (Power *et al.*, 2010). This allows the entire network topology to be summarized in a simple connectivity matrix, where the rows and columns correspond to the different nodes, and each cell contains their pairwise connectivity weight (Figure 6.7). This high degree of data compression is one important reason why graph theory measures have become especially useful summaries for the largest and most complexly connected systems like Facebook social networks or the human brain. From these raw connectivity matrices, one can derive a variety of summary meta-statistics describing different aspects of their "global" connectivity across all nodes. Similarly, the connectivity of individual nodes can be summarized through a variety of "nodal" measures. As a powerful qualitative visualization of the network topology, the layout from the raw connectivity matrix can be displayed as a series of interconnected nodes and edges, and simultaneously encoded with any of these graph theoretical measures (Figure 6.8).

6.4.1 UCLA Multimodal Connectivity Database (UMCD)

Since initially-large and diverse brain imaging data sets are highly distilled for graph-based approaches into a set of uniform input connectivity matrices, these graph theory techniques lend themselves especially well to fully-automated processing pipelines. One such effort is the UCLA Multimodal Connectivity Database (UMCD), which provides a public web-based interactive interface to an automated graph theory pipeline, and outputs many of the most common global and nodal connectivity measures.

To automate the submission and retrieval of UMCD processing requests, we authored R-UMCD, an R interface to the web API that powers this online tool. With this wrapper, a batch of results for a list of subjects can be compiled directly from within an R session, and then immediately analyzed and visualized.

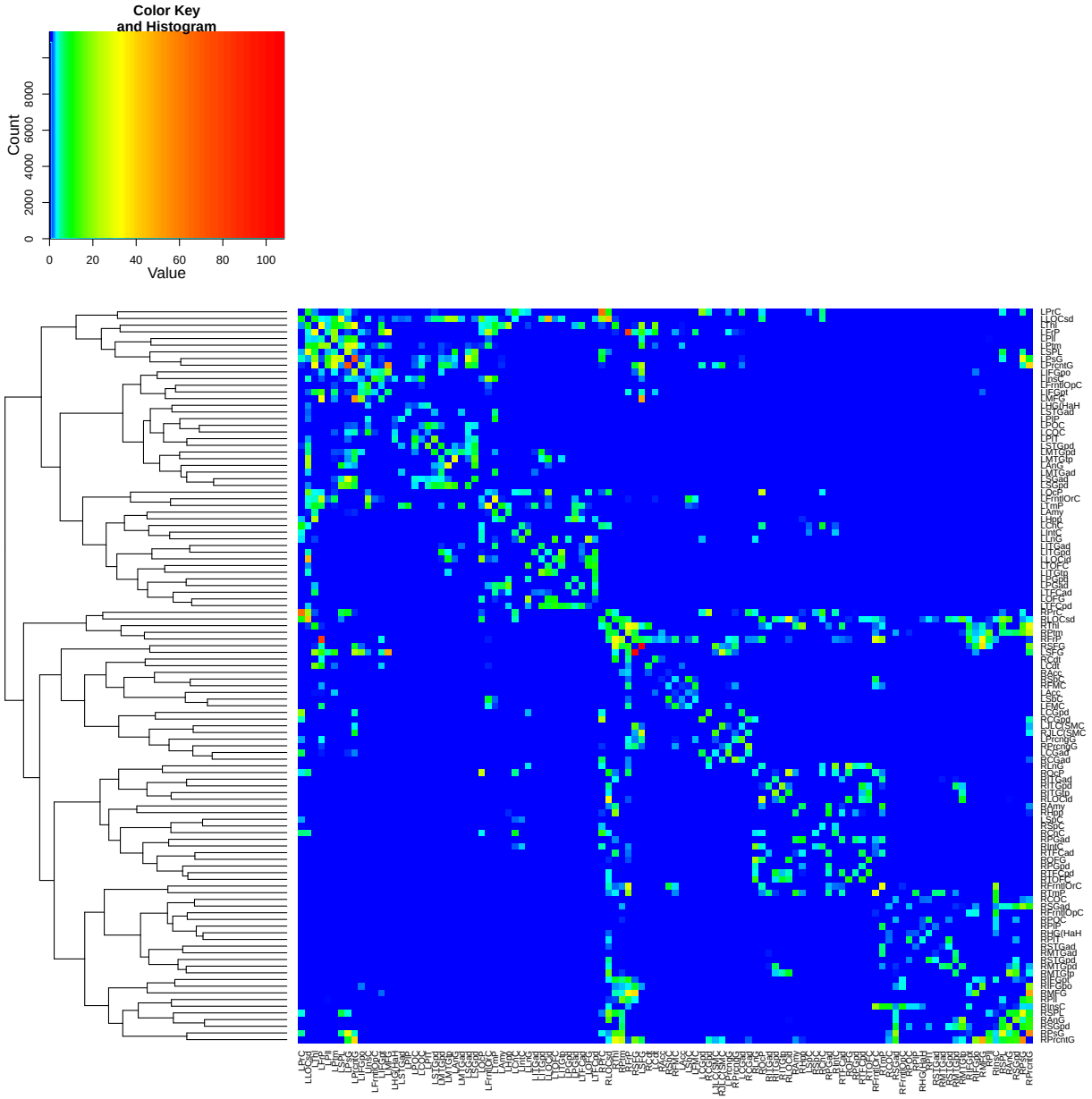


Figure 6.7: Study-wide DTI connectivity matrix ($n = 51$). Color indicates average fiber count. Nodes clustered via hierarchical agglomerative algorithm.

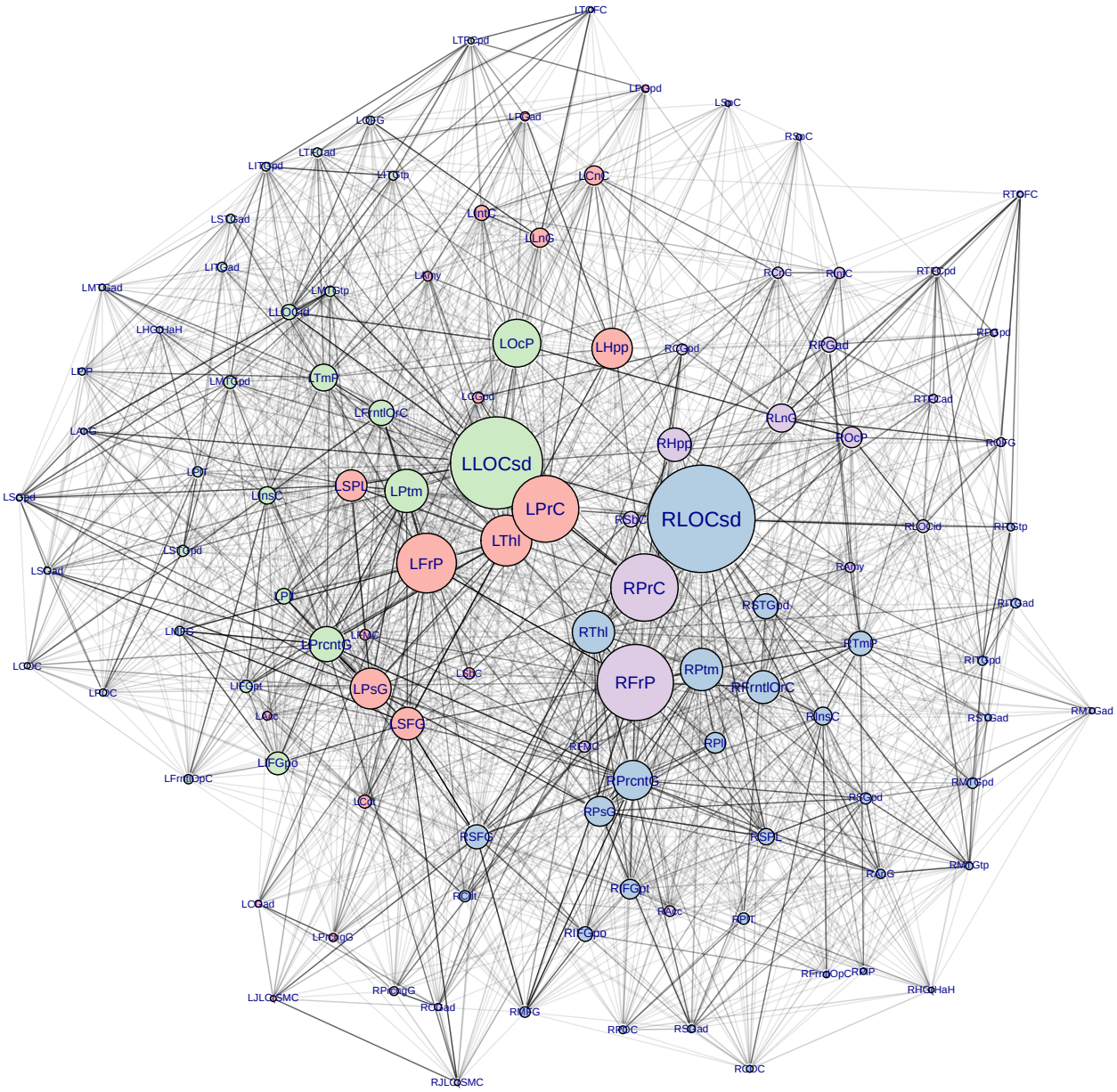


Figure 6.8: Study-wide graph visualization. Edge transparency indicates average fiber count between nodes. Node size indicates betweenness centrality. Node color indicates eigenvector community membership.

6.4.2 Results

Graph theoretical measures of white matter network connectivity were examined in all 51 subjects. Streamline tractography fiber counts were calculated in a pairwise fashion between 110 cortical and subcortical ROIs derived from the Harvard-Oxford atlas, and used to generate white matter connectivity matrices for each individual. These were aggregated by group and used to generate group-mean network metric summaries via the UMCD. Qualitatively, the FASD group mean exhibited lower overall connection density, and, when both groups were normalized to a prescribed sparsity of 20%, lower modularity and lower normalized clustering coefficient. Overall, this pattern leads to a lower sigma measure of "small-worldness" in the FASD group (Table 6.3).

	Measure	control_mean	fasd_mean
1	Characteristic Path Length	2.00	1.99
2	Clustering Coefficient	0.58	0.57
3	Gamma	1.09	1.09
4	Global Efficiency	0.57	0.57
5	Lambda	2.09	1.99
6	Modularity (Q)	0.38	0.36
7	Number of Components	1.00	1.00
8	Raw Density (%)	30.46	29.24
9	Sigma	1.92	1.83

Table 6.3: Qualitative comparison of global graph theory measures between group mean networks.

To see if this qualitative patten was statistically significant, a full between-group analysis was conducted using global measures derived from single-subject connectivity matrices. Age, group, and age:group interactive effects were considered in a linear model that was fit for each global measure. Even without considering the large number of comparisons, all of the terms for all of the metrics are nearly at or below threshold (Table 6.4). The lack of any prominent relationship is also clear when examining plots of these data (Figure 6.9).

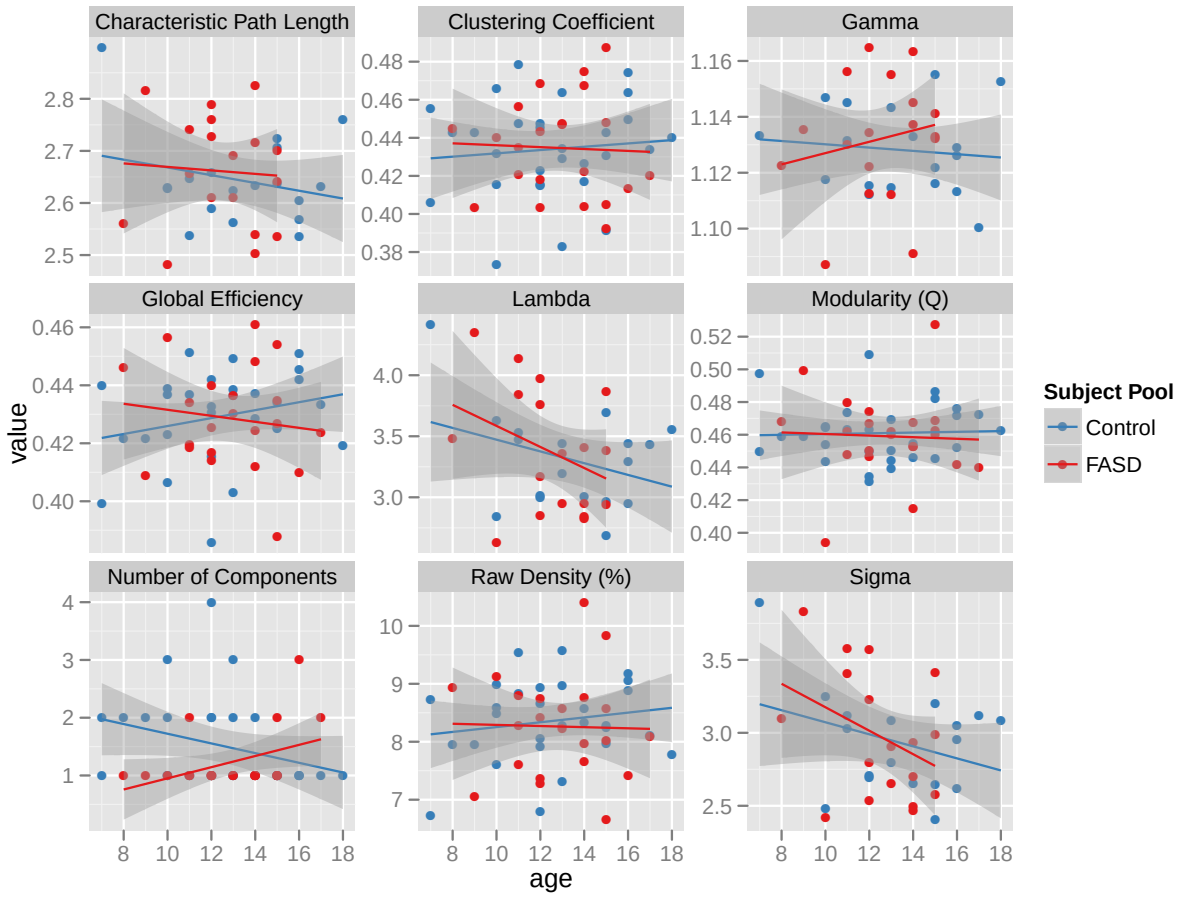


Figure 6.9: Global graph theory measures by age and group

	Measure	Term	Estimate	Std. Error	t-value	Pr(>|t|)
1	Characteristic Path Length	age	-0.01	0.01	-0.89	0.38
2	Characteristic Path Length	‘Subject Pool‘FASD	-0.04	0.19	-0.21	0.83
3	Characteristic Path Length	age:‘Subject Pool‘FASD	0.00	0.01	0.29	0.78
4	Clustering Coefficient	age	0.00	0.00	0.50	0.62
5	Clustering Coefficient	‘Subject Pool‘FASD	0.02	0.04	0.45	0.65
6	Clustering Coefficient	age:‘Subject Pool‘FASD	-0.00	0.00	-0.45	0.66
7	Gamma	age	-0.00	0.00	-0.31	0.76
8	Gamma	‘Subject Pool‘FASD	-0.03	0.04	-0.94	0.36
9	Gamma	age:‘Subject Pool‘FASD	0.00	0.00	0.99	0.33
10	Global Efficiency	age	0.00	0.00	1.26	0.21
11	Global Efficiency	‘Subject Pool‘FASD	0.03	0.02	1.19	0.24
12	Global Efficiency	age:‘Subject Pool‘FASD	-0.00	0.00	-1.25	0.22
13	Lambda	age	0.03	0.03	0.90	0.37
14	Lambda	‘Subject Pool‘FASD	0.47	0.76	0.62	0.54
15	Lambda	age:‘Subject Pool‘FASD	-0.02	0.06	-0.35	0.73
16	Modularity (Q)	age	0.00	0.00	0.15	0.88
17	Modularity (Q)	‘Subject Pool‘FASD	0.01	0.03	0.21	0.83
18	Modularity (Q)	age:‘Subject Pool‘FASD	-0.00	0.00	-0.28	0.78
19	Number of Components	age	-0.08	0.04	-1.95	0.06
20	Number of Components	‘Subject Pool‘FASD	-2.58	0.99	-2.61	0.01
21	Number of Components	age:‘Subject Pool‘FASD	0.18	0.08	2.37	0.02
22	Raw Density (%)	age	0.04	0.05	0.80	0.43
23	Raw Density (%)	‘Subject Pool‘FASD	0.55	1.20	0.46	0.65
24	Raw Density (%)	age:‘Subject Pool‘FASD	-0.05	0.09	-0.56	0.58
25	Sigma	age	0.03	0.03	0.99	0.33
26	Sigma	‘Subject Pool‘FASD	0.49	0.63	0.78	0.44
27	Sigma	age:‘Subject Pool‘FASD	-0.02	0.05	-0.50	0.62

Table 6.4: Between-group analysis of global measures

6.5 Discussion

6.5.1 Age effects

On the whole, this investigation was able to successfully demonstrate the expected patterns of developmental maturation in the white matter — including decreasing diffusivity, and increasing fractional anisotropy. However, the robustness of these findings, particularly for FA, was lower than expected. For instance, while we expected broadly increasing FA across many brain regions, significant effects were isolated to only modest regions of the brainstem, genu of the corpus callo-

sum, and right cingulum on the voxelwise TBSS analysis (Figure 6.2A). The pattern was closer to what was expected when the analysis was repeated while excluding the methamphetamine-exposed subjects (Figure 6.3). This makes sense generally because a uniform typically-developing control group is expected to be one of the "cleanest" samples for demonstrating simple age effects. With the addition of patient groups, even when modeling different intercepts, the increased noise in the sample could outweigh the expected increase in power due to the higher *n*. This appears to be the case here. The confound of methamphetamine exposure is especially relevant, since we have shown this to be associated with effects on FA (Chapter 2) that could further muddy the usefulness of the FASD group for demonstrating age effects. The lateralization of the arcuate fasciculus, in terms of the number of streamlines in each tract (Lebel and Beaulieu, 2009), is another typically-large developmental effect that we were able to demonstrate only weakly.

6.5.2 FASD effects

When correcting for multiple comparisons across all spatial positions, neither the voxelwise TBSS analysis, nor the along-tract tractography analysis, were able to demonstrate any significant group effects between the FASD and Control subjects. The graph theory network-based approach was even less fruitful. This is not surprising given the low *n* and contamination in the FASD group. However, when more omnibus tests were performed on the along-tract tractography data, significant position:group interactions in predicting FA were observed in several regions (§ 6.3.2). The *t* tests along these specific tracts showed these effects generally to be due to decreased FA in the FASD group, which is consistent with previous literature showing broadly decreased FA and increased MD in these individuals (Lebel *et al.*, 2008a; Norman *et al.*, 2009). The bilateral nature and similarity in along-tract localization of the findings in the inferior frontal occipital fasciculus (IFO) and corticospinal tract (CST) lend somewhat more confidence to these observations. The posterior IFO localization, especially, is consistent with a previous voxelwise TBSS analysis that also demonstrated lower FA in the same region bilaterally (Fryer *et al.*, 2009). Abnormalities in this region are also functionally consistent with deficits in visuospatial processing seen in FASDs

(Mattson and Riley, 1998; Mattson *et al.*, 1998). However, it is conspicuous then that we failed to observe *any* effects in the corpus callosum, which typically shows the largest abnormalities, and has been broadly replicated across studies (Wozniak and Muetzel, 2011). Finally, looking at across-tract properties, the fact that the FASD group had shorter tracts overall, and a trend toward more tracking failures, was expected based on the global microcephaly that is commonly seen in these individuals (Spohr *et al.*, 2007).

6.5.3 Voxelwise analysis vs. tractography

The voxelwise TBSS analysis §6.2 and the streamline tractography analysis §6.3 take very different approaches to data processing, and yet both get used with the goal of demonstrating the same ultimate phenomena. So which is better? Although there are many differences between the two approaches, one of the most important is that TBSS implements a nonlinear warp on the scalar FA maps to bring different subjects into register, while tractography takes a more implicit approach to registration by imposing anatomical constraints on the vector diffusion directionality fields. This means that the TBSS registration will be driven by the high-contrast edges between white/gray matter, and other regions of high/low fiber coherence on the FA maps, but will be relatively blind to important edges in directionality space (*e.g.* two highly coherent bundles passing adjacent to each other). Additionally, because the anatomical constraint in voxelwise registrations is that all brains should have the same general morphology, they are less accommodating to large structural differences that can be seen in disease or even across the span of typical development. On the other hand, there are certain regions of the white matter that typically do not segregate into any of the standard white matter tract groups with tractography, and thus are never even analyzed with this approach. This is especially true with DTI-type analyses that only consider a single intravoxel fiber population. In these analyses, regions of crossing fibers act to cleave off the minor intravoxel fiber populations, and leave these fiber groups isolated from the rest of the network. One classic example of this is the 3-way intersection of the corticospinal tract, superior longitudinal fasciculus, and lateral corpus callosum extension, where the corticospinal tract dissections almost never include

projections to the more lateral aspects of the primary motor strip.

6.5.4 Participants

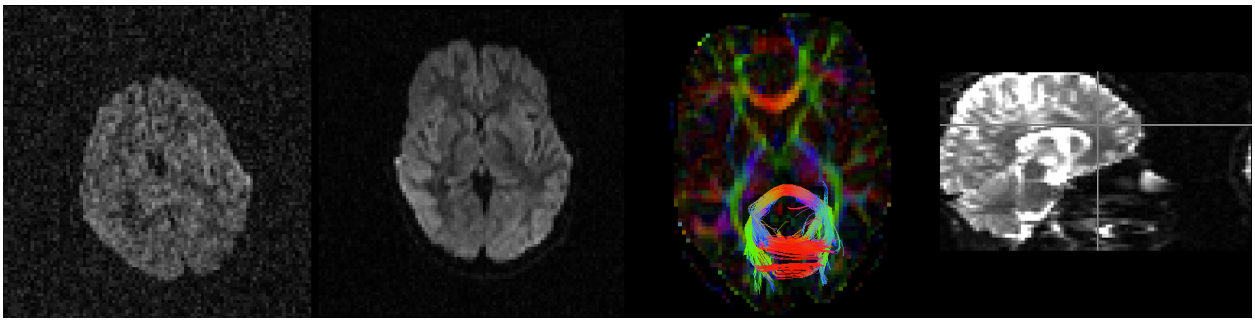
The sample characteristics of this study are clearly its most important limitation. When data collection started, the goal was to collect 70 subjects in each group. Actual data collection fell far short of this mark, with only 70 subjects *total* enrolled in the study. Further, if the subjects with methamphetamine exposure are excluded — which is a reasonable thing to do based on potential confounds — only 12 subjects remain in the FASD patient group. With baseline power for detecting group effects this low, it is no surprise that we were unable to demonstrate the same level of widespread white matter abnormalities that have been previously reported (Wozniak and Muetzel, 2011). This low *n*, especially in the FASD group, also prevents a thorough investigation of the more complex aspects of our initial proposal, which concerned mapping the nonlinear developmental trajectories (See Aim 1b) and machine learning classification of FASD subjects from controls (See Aim 2d).

6.5.5 Data quality

Upon reflection, the overall image quality of this data set is also likely contributing to lower power and our failing to demonstrate specific hypothesized findings. During processing of this data set, several data quality issues were noted — all of which can be generally grouped under the umbrella of "growing pains" as research was transitioned to a new Siemens Trio scanner at the UCLA Brain Mapping Center in 2008. These problems reduced the overall yield of usable scans, and, for those scans that *were* included, are likely still contributing in subtler ways.

Low SNR: For reasons unknown to us, the overall SNR in the raw data fluctuated during the initial months of operation, causing some scans to have unacceptably low SNR. Compare Figure 6.10A to Figure 6.10B.

Artifacts: Focal artifacts were also a problem. By far the most severe was a vibration artifact that would manifest prominently in the processed data as spurious fibers along the x axis



(A) Low SNR **(B)** Typical SNR **(C)** Vibration artifact **(D)** Clipping/aliasing

Figure 6.10: Data quality issues

(Figure 6.10C).

New protocol: Due to the new acquisition protocol, with a new field of view and total z slab thickness, subject positioning issues became a common problem. This included aliasing along the A–P phase-encode direction, clipping of the brain either superiorly or inferiorly along the z axis, and sometimes even both (Figure 6.10D).

6.5.6 Conclusion

In spite of the limitations of this sample, we were able to successfully demonstrate the more general aspects of developmental maturation in DTI metrics across groups. Although underpowered, patterns of abnormalities in fractional anisotropy — a DTI metric sensitive to myelination — were also demonstrated in the FASD group, and are broadly consistent with previous voxelwise and tract-based results. Further, the application of along-tract analysis techniques has provided some promising insight, albeit preliminary, into the within-tract localization of these findings.

CHAPTER 7

Machine learning classification of attention deficit hyperactivity disorder ($n = 973$)

7.1 Introduction

As part of our NRSA grant proposal to study the effects of prenatal alcohol exposure on white matter microstructure and its relationship to executive function (Chapter 3), we had originally intended to investigate the usefulness of neuroimaging features (*e.g.* along-tract DTI measures as in § 6.3) for the classification of individuals with FASDs from typically-developing controls. In light of the sample size limitations of the UCLA FASD data set, we instead decided to shift our focus to a phenotypically-related neurodevelopmental disorder — Attention Deficit Hyperactivity Disorder (ADHD). This was made possible through the recent public release of the exciting 8-site ADHD–200 data set, and a collaborative team effort between myself, Jesse Brown, Jeffrey Rudie, and Pamela Douglas (from UCLA), and Zarrar Shehzad (from Yale University). Although not specific to FASDs, the hope for me was that this effort would also reveal broader insights into the diagnostic classification potential of neuroimaging features that could be generalizable *across* different neurodevelopmental disorders.

ADHD is an important neurodevelopmental disorder that is estimated to affect approximately 10% of children (Froehlich *et al.*, 2007; Larson *et al.*, 2011; Merikangas *et al.*, 2010), with an estimated economic cost to society in the tens of billions of dollars per year (Pelham *et al.*, 2007). Overall, ADHD is more common in males than females, with over a 2:1 ratio ("MMWR" 2010). Diagnosis is made based on behavioral symptoms reported by the child and parent, in areas related

to inattention, impulsivity, and hyperactivity. The DSM-IV recognizes three ADHD subtypes, predominantly hyperactive-impulsive (ADHD-H), predominantly inattentive (ADHD-I), and combined hyperactive-impulsive and inattentive (ADHD-C). Like many Axis I disorders, diagnosis also hinges on the degree to which these impairments actually interfere with daily life at school, home, and/or work (*DSM-IV* 2000). Treatment includes medication, behavioral therapy, and/or educational interventions (Wolraich *et al.*, 2011). Although ADHD is clinically-diagnosed disorder, there have also been a variety of neuroimaging findings reported in the literature. These include altered functional activation and structural connectivity (Liston *et al.*, 2011), and cortical development (Shaw *et al.*, 2007; Shaw *et al.*, 2006b). Importantly, while FASDs are driven by a different upstream insult — prenatal alcohol exposure — they share many behavioral phenotypic similarities with children with ADHD (Fryer *et al.*, 2007).

The ADHD–200 initiative was structured into a machine learning contest — The ADHD–200 Global Competition — in order to spur interest in the project across disciplines (neuroscience, statistics, computer science, etc.). The broad goal was to investigate whether these previously-observed neuroimaging biomarkers of the disorder could be translated into a classifier that could be able to accurately classify new individuals. A training data set was released first, and included imaging data from 776 individuals (491 typically-developing (TD) and 285 ADHD), their diagnostic class labels (TD or ADHD subtype), and accompanying demographic information (*e.g.* site, age, IQ). Imaging data for all subjects included one or more resting-state functional MRI (fMRI) scans, and a high resolution T1-weighted anatomical scan. This portion of the data was used to design and train an imaging-based diagnostic classifier that could then be used to predict the unknown group memberships of additional subjects based on their imaging and demographic features alone. The different classifiers entered into the contest were then judged based on their diagnostic prediction accuracies across a final hold-out (*i.e.* previously-unreleased) test data set. A greater penalty was placed on false-positive (type I) errors, which is relatable to many "confirmatory"-type diagnostic tests in the real world where there is some risk of harm associated with the treatment itself and therefore false positives must be avoided.

7.2 Demographics

7.2.1 Numbers of subjects by diagnosis and site

Even before diving into the more complex imaging features, the demographics data were able to provide extremely valuable insight. The most basic finding is that the prevalence of ADHD across the *entire* training set is 37%. Ignoring the two sites without any ADHD subjects in their training data, this number jumps up to 46%. This simple fact is critical to appreciate because — assuming the training set is representative of the final test set (See §7.2.4) — it lays down the general pretest probability for whatever final classifier is developed. For example, given a test subject with equivocal imaging features, we would like a classifer that does not simply flip a coin and assign their diagnosis based on 50/50 chance, but would favor a moderately higher probability diagnosis of TD based on the higher representation of TD subjects across the training set. In diagnostic testing terms, the lower the pretest probability of ADHD, the lower the positive predictive value of our test. After all, even if we develop a test that is 99.99% accurate at diagnosing some disease, if we know for a fact that the prevalence of the disease in our population is 0%, than *all* of the positive test results are still going to be false positive type I errors. The greatest class prevalence in the training data set is also called the *no-information rate*, and is a more useful benchmark of what we need to beat in order for our classifier to be useful. For example, if our classifier is 80% accurate at diagnosing an individual as ADHD or TD, but we know that 9 out of 10 people in the real world are TD, then we are better off throwing away our classifier and just calling everyone TD! Consequently, the barrier to developing a useful diagnostic test is higher in the real world, where the disease prevalences are typically lower than in controlled studies, and there are additional factors to consider such as cost and potential treatment risks. Consider autism, where the low real-world prevalence of around 1% generally precludes the use of any imaging-based classifiers in a screening/diagnostic context, and instead limits their usefulness to a more descriptive role in research studies where the pretest probability is already high. This exact situation put the authors of an otherwise well-written paper in hot water when their later comments combined with an over-zealous press release to overstate

the implications of their work (Ecker *et al.*, 2010; Heneghan, 2010).

	TD	ADHD-C	ADHD-H	ADHD-I	**Sum**
Peking	116	29	0	49	194
KKI	61	16	1	5	83
NI	23	18	6	1	48
NYU	99	77	2	44	222
OHSU	42	23	2	12	79
Pitt	89	0	0	0	89
WashU	61	0	0	0	61
Sum	491	163	11	111	776

Table 7.1: Number of subjects in training set data, by site and diagnosis

	TD	ADHD
Peking	0.60	0.40
KKI	0.73	0.27
NI	0.48	0.52
NYU	0.45	0.55
OHSU	0.53	0.47
Pitt	1.00	0.00
WashU	1.00	0.00

Table 7.2: Overall class proportions by site. ADHD-H subjects excluded (no test set data).

Breaking down these numbers further, we can see that there are also prominent differences in ADHD prevalence, as well as ADHD subtype ratios, between sites (Table 7.1). The ADHD-H is the lowest represented among the three subtypes, and we were told that no subjects with this diagnosis would be included in the test set. Therefore these subjects were excluded up front. Of the remaining subjects, the highest prevalence of ADHD is 55% at the NYU site, and the lowest is 27% at the KKI site (Table 7.2). Of the two remaining subtypes, ratios varied widely between sites — from 18:1 ADHD-C:ADHD-I at the NeuroIMAGE site to 29:49 at the Peking site. While these specific numbers are of course artifacts of sampling bias across the individually-designed studies that joined ADHD–200, they again highlight the importance of tuning our classifier to the variable prevalence within sub-groups in our population. Similar diversity also exists in the real world. Perhaps some regions of the country have a higher prevalence of a certain disease due to differing demographic,

genetic, or environmental factors? Or maybe care providers at one institution are more in tune to making a certain diagnosis because a specialized clinic exists for it at their site and keeps it fresh in their minds? All of these things must be considered. Shifting back to the ADHD–200 sample, we continued to build our intuitive understanding of how we wanted our classifier to perform. We already knew that designing a classifier to resolve a behaviorally-diagnosed disease like ADHD from TD, based on brain imaging data alone, would be challenging. On top of that, it seemed unlikely that imaging features would outperform these strong baseline expectations about which subtype to expect at which site. Therefore, we decided early on that we would concentrate our effort on classifying ADHD from TD, generally, and would default to these prior expectations for assigning ADHD subtypes.

7.2.2 Gender

The next way to slice the data is by gender (Table 7.3). In the general population, ADHD is far more common in males than in females, so we expected this feature would be very useful. The first thing to notice is that, in aggregate across the training set, ADHD is far less common in females than in males. Ignoring the sites with no ADHD subjects, the prevalence of ADHD in females is 27%, but in males it is 54%. This effect is huge, and immediately told us that our pretest expectations about ADHD diagnosis would vary prominently based on whether a given test set subject is male or female. Similar to the ADHD prevalence across gender, these numbers also vary strongly by site. For example, at Peking, only 13% of female training subjects are ADHD, while at NYU 64% of males are ADHD.

7.2.3 Age and IQ

Beyond looking at prevalence rates across sites and genders, the two main *continuous*-valued demographic features made available were age and a full-scale IQ score. The relationships between age, IQ, and diagnosis — together with how these vary by site and gender — can all be very effec-

	TD	ADHD-C	ADHD-I	**Sum**
Peking	45 (86.5)	0 (0.0)	7 (13.5)	52 (100.0)
KKI	27 (73.0)	9 (24.3)	1 (2.7)	37 (100.0)
NI	12 (75.0)	4 (25.0)	0 (0.0)	16 (100.0)
NYU	52 (65.8)	12 (15.2)	15 (19.0)	79 (100.0)
OHSU	24 (70.6)	4 (11.8)	6 (17.6)	34 (100.0)
Sum	160 (73.4)	29 (13.3)	29 (13.3)	218 (100.0)

(A) Females

	TD	ADHD-C	ADHD-I	**Sum**
Peking	71 (50.0)	29 (20.4)	42 (29.6)	142 (100.0)
KKI	34 (75.6)	7 (15.6)	4 (8.9)	45 (100.0)
NI	11 (42.3)	14 (53.8)	1 (3.8)	26 (100.0)
NYU	47 (33.6)	64 (45.7)	29 (20.7)	140 (100.0)
OHSU	18 (41.9)	19 (44.2)	6 (14.0)	43 (100.0)
Sum	181 (45.7)	133 (33.6)	82 (20.7)	396 (100.0)

(B) Males

Table 7.3: Numbers of subjects by site, diagnosis, and gender. Percentages of row totals are shown in parentheses. ADHD-H subjects excluded (no test set data). Pittsburgh and Washington University sites excluded (no ADHD subjects).

tively visualized in a single display (Figure 7.1). There are several important things to notice here. First, considering the marginal distributions, there is no striking gender effect on age or IQ (*i.e.* age and IQ are relatively well-matched across genders), but there is a strong site effect on both age and IQ. At the extremes, OHSU doesn't have *any* subjects older than 12 years old, and Pittsburgh doesn't have any subjects younger than 10. Similarly, OHSU females have exceptionally high IQs. Secondly, considering main effects, there is a strong correlation between lower IQ and ADHD diagnosis, but no appreciable age effect is present. Lastly, considering the joint usefulness of age and IQ for predicting diagnosis, we can see a large degree of variability between sites and genders. For example, whereas Peking subjects separate nicely based on these features, NYU subjects do not.

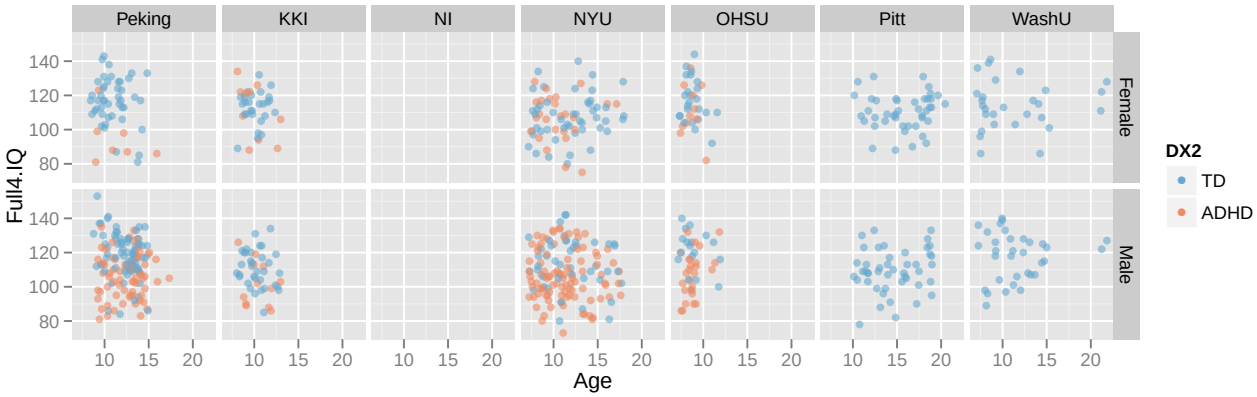


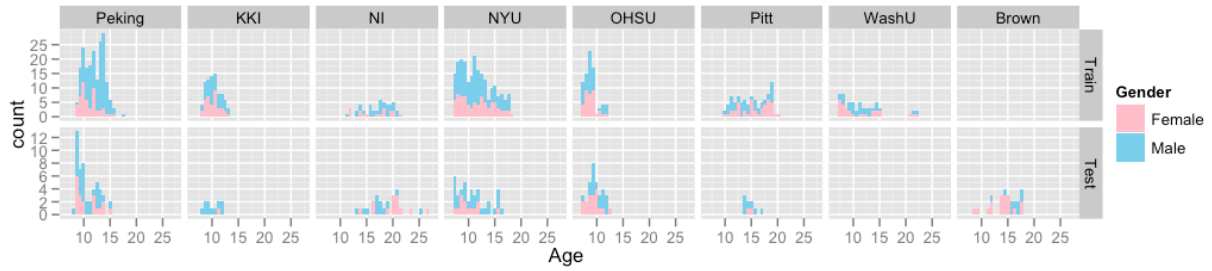
Figure 7.1: IQ vs. age, by diagnosis, site, and gender

7.2.4 Training set vs. test set

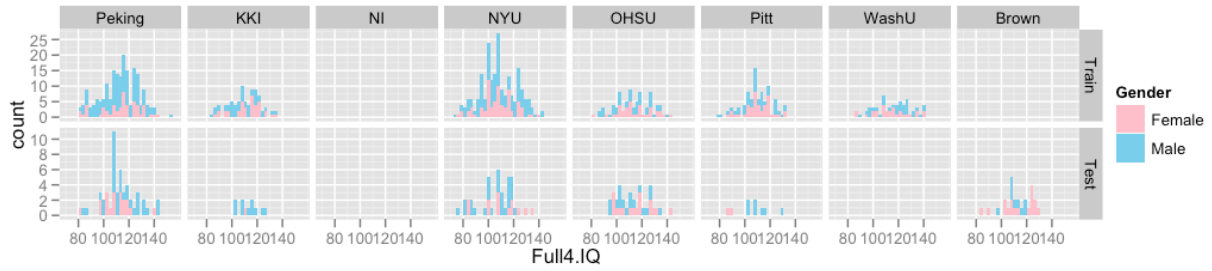
One of our most important discussions was whether we could expect similar demographic/prevalence biases in the test set. In the real world we would expect that we could, but, in the context of a designed study, we were unsure. Perhaps, specifically to emphasize the classification potential of *imaging* features, the test set was age and IQ matched between ADHD and TD, and hand-picked to be 50:50 ADHD:TD. To investigate this, we compared the test set to the training set on the features that we were given — namely age and IQ (Figure 7.2). We reasoned that comparing these distributions would give us insight into whether or not we might suspect something-other-than-random sampling in the test set. If this were the case, we would consider abandoning our plan to make these features central to our classifier. For example, considering the large IQ difference between male Peking ADHD and TD subjects (bottom left panel in Figure 7.1), we might suspect that the test set was being designed to match on IQ if the IQ distribution was generally lower than in the training set. From examining Figure Figure 7.2, however, it does appear that the training set is generally representative of the test set.

7.2.5 Site-specific observations

Several important site-specific aspects of the ADHD–200 dataset also become apparent when examining the demographics features. Most strikingly, the Washington University site is not represented



(A) Age



(B) IQ

Figure 7.2: Comparison of training/test sets, by site and gender

in the test data set, and the Brown site is not represented in the training data set. A third site, Pittsburgh, has both training data and test data, but no ADHD representation in the training data. Given the large site effects that have already been discussed in some detail (and will be discussed further, with respect to imaging features, in §7.3), these missing data pose several critical questions:

- Should the **Washington University** data even be included when training our classifier, or ignored completely?
- How can we utilize the **Pittsburgh** training data without inaccurately biasing the classifier towards a TD diagnosis at this site?
- Finally, how can we most effectively accommodate test subjects from the **Brown** site, given that there are no training data available?

```
Confusion Matrix and Statistics

          Reference
Prediction   TD ADHD-C ADHD-I
    TD     2720    773    631
    ADHD-C  287    571    212
    ADHD-I   93     96    227

Overall Statistics

               Accuracy : 0.6271
                 95% CI : (0.6143, 0.6398)
    No Information Rate : 0.5526
    P-Value [Acc > NIR] : < 2.2e-16

                  Kappa : 0.2973
 Mcnemar's Test P-Value : < 2.2e-16

Statistics by Class:

                     Class: TD Class: ADHD-C Class: ADHD-I
Sensitivity             0.8774        0.3965       0.21215
Specificity             0.4406        0.8803       0.95837
Pos Pred Value          0.6596        0.5336       0.54567
Neg Pred Value          0.7443        0.8086       0.83770
Prevalence              0.5526        0.2567       0.19073
Detection Rate          0.4848        0.1018       0.04046
Detection Prevalence    0.7351        0.1907       0.07415
```

Figure 7.3: Classifier statistics for 3-class RBF SVM trained on demographics

7.2.6 Classification based on demographics data alone

As a baseline reference for building our imaging-based classifier, we explored the generalization performance we could expect from training a classifier on only the demographics features. Age, gender, site, and IQ features were included from the four sites with complete data (Peking, KKI, NYU, and OHSU), and used to train an RBF SVM classifier (See §7.4.2 for an explanation). Using these features alone, predicted generalization accuracy was 62.7%. The no-information rate for this subset of the data was 55.3%, suggesting that demographics features alone are able to moderately improve classification performance Figure 7.3. Taking into account the fact that greater emphasis was placed on correct TD diagnoses, this would correspond to us achieving 65.5% of the total possible points.

7.2.7 Summary

On the whole, this analysis of the demographics data revealed a very strong feature set, with prominent site and gender effects. It was clear to us at this early stage that these demographics features would form the core of our classifier, and would set the bar high for finding imaging features that could further clarify the picture.

7.3 Feature pool

7.3.1 Imaging

7.3.1.1 Structural

Under an effort led by Jesse Brown, T1-weighted anatomical MRI scans were processed with the Freesurfer processing pipeline (Fischl and Dale, 2000). This generates segmentations of white matter, gray matter, and subcortical volumes. Cortical areas are also used to generate a mesh model of the cortical surface, which is then subdivided into different cortical regions (*e.g.* precentral gyrus, superior frontal gyrus, pars triangularis, etc.). For each, features are generated for cortical surface area, thickness, volume, and local curvature.

7.3.1.2 Functional

Resting-state fMRI scans were preprocessed with standard AFNI and FSL tools by other members of the community (the "Neuro Bureau"), and made available to all users. Under an effort led by Jeffrey Rudie, for each subject the pairwise average time series correlations were calculated between all sets of nodes (derived from an atlas containing different cortical regions of interest), and used to populate a connectivity matrix (for example, Figure 6.7). Several atlases were explored, ranging from the Harvard–Oxford atlas (a relatively course anatomical atlas with around 100 nodes), all the way up to the CC400 atlas (a fine functionally-derived atlas with 400 nodes). As one option,

these raw times series correlations were used directly as features. As another, they were used to generate nodal and global graph theoretical summaries of the network properties, and *these* were used as features. Under an effort led by Pamela Douglas, independent component analysis (ICA) was performed, and the number of independent components needed to describe a certain percentage of the total variance in the data was used as a feature. Finally, under an effort led by Zarrar Shehzad, several emerging techniques for functional brain network analysis were also used to generate complementary sets of features. These included node-wise frequency analysis (the bins from a frequency-domain power spectrum were used as features), and regional homogeneity analysis (ReHO; in which the similarity is calculated between the time series at any given point in the brain, and those of its surrounding neighbors) (Zang *et al.*, 2004).

7.3.2 Other ideas

In addition to the imaging and demographics features, we also investigated several other more creative features.

7.3.2.1 QC results

Based in part on our experience scanning children with FASDs, the majority of whom also have attention or hyperactivity problems, we reasoned that the quality of the MRI scans could provide a useful feature for classifying ADHD from TD. We examined two scores: 1) the total number of fMRI scans performed, and 2) the fraction of these that were tagged as "usable" in the demographics table. As expected, ADHD subjects had a significantly higher number of scans performed ($P = 0.0005$), but had a significantly lower fraction of usable scans ($P = 0.0003$). In other words, ADHD subjects took more tries to get the required data needed for the study.

7.3.2.2 Motion

Standard fMRI processing includes registration of the individual time series volumes, and this step outputs the translations and rotations that were needed to bring each volume into alignment. This set of 6 rigid-body registration parameters reflects subject motion between two adjacent time points. We examined the distributions of the means, as well as the maximums, of these parameters between ADHD and TD (Figure 7.4). As expected, we found highly significant group differences.

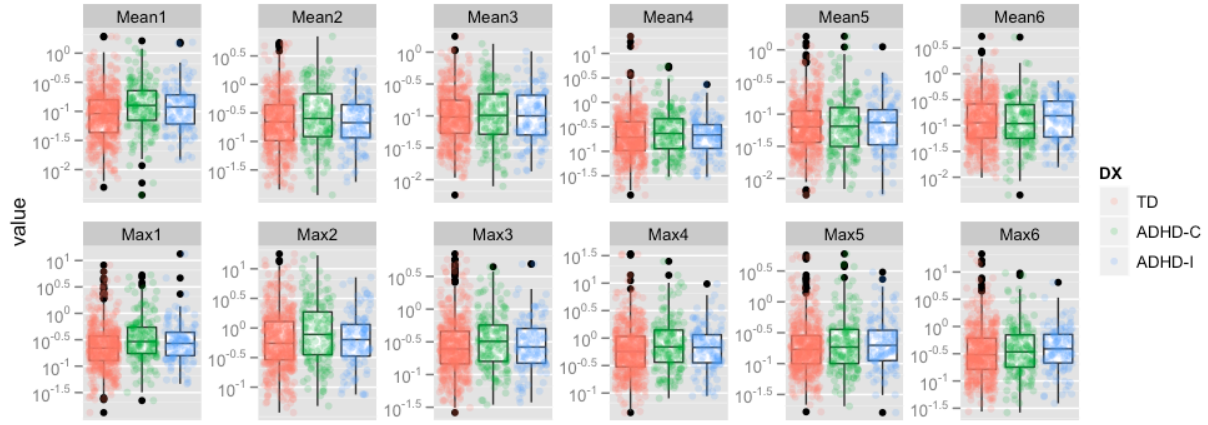


Figure 7.4: Motion parameters from fMRI data registration

7.4 Feature selection

The field of neuroimaging is not alone in facing the challenge of an accelerating explosion of data. Whether it be genetics and genome-wide sequencing data, medicine and advancing electronic medical records keeping, or commercial technology companies and their minute-by-minute user activity data, diverse fields are all converging on the same major challenge — How do you distill these large, multimodal, and multivariate data sets down into useful summaries that can be actionable by their human wranglers?

7.4.1 Feature ranking with SVM–RFE

If all of the many tens of thousands of features described in §7.3 are included in the same classifier, the individual variability across the many useless features will typically overwhelm the discriminating power of the subset of features that are actually useful. Therefore, in many classification applications, the first step is some form of dimensionality reduction or feature selection. We chose to apply the support vector machine recursive feature elimination (SVM–RFE) algorithm (Guyon *et al.*, 2002) because of its established history and successful application to microarray-based diagnostic classification (Johannes *et al.*, 2010; Shi *et al.*, 2011) — a similarly medical and connectivity/network type of application.

SVM–RFE, as its name would suggest, works backwards from the initial full set of features and eliminates the least "useful" feature on each recursive pass. In contrast to optimization methods that can revisit locations in feature space (*e.g.* genetic/evolutionary algorithms (Vafaie and Imam, 1994)), it is important to note that this is a *greedy* approach in that once a feature is cut, it is cut for good. The criterion used to judge feature usefulness in SVM–RFE is the feature weight from a linear support vector machine (SVM) fit to the data set. Linear SVM is a linear discriminant, in that it seeks to find a *linear* combination of the features that allows for the best classification of groups. Whereas the classical Linear Discriminant Analysis (LDA) interpretation seeks to maximize the ratio of the between-class variance to the within-class variance in the standard ANOVA sense, SVM seeks a discriminant function that maximizes the distance (*i.e.* margin) to the nearest training set observation of either class. The theory was originally described by Vapnik and Lerner (1963), and later extended to accommodate the exceedingly common situation where the classes are *not* completely separable and therefore some training examples must remain mislabeled in the solution (Cortes and Vapnik, 1995). This decision boundary ends up as a line in two dimensional feature space and as a higher dimensional hyperplane when more features are present. Because linear SVM assigns weights multivariately to all remaining features at once, it has the ability to accommodate highly correlated features, as well as potential mutual information between features that might not be very useful on their own. This approach contrasts with univariate correlation-based feature

ranking, where features are ordered, for example, by conducting simple between-group *t*-tests for each (Guyon *et al.*, 2002).

In practice, the general idea is very intuitive. Consider the simplified 2-dimensional, 2-class, case of distinguishing ADHD from TD subjects from the Peking site based on age and IQ alone. This corresponds to the bottom-left panel in Figure 7.1. If we plot the results of the linear SVM fit, we can see that the decision boundary cuts more along the IQ axis. This means that the IQ feature has a higher weight than age, and that age would be dropped first in the recursive elimination algorithm.

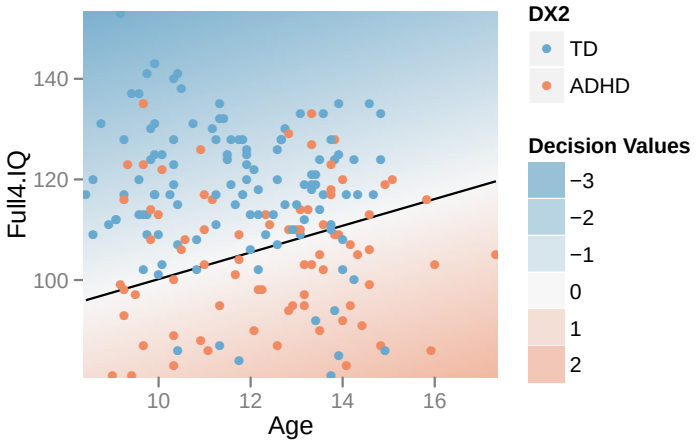


Figure 7.5: Linear SVM fit to Peking age and IQ data

Wrapped around the basic SVM–RFE algorithm, we also included two enhancements that were very useful for our application:

CV-based stabilization of feature rankings: Due in part to its multivariate nature, and also simple sampling variability, some of the feature rankings that get output from the SVM–RFE algorithm can be unstable. Take, for example, the case of two highly correlated — but useful — features. Their joint weight will initially be split between the two features. When the one with the (arbitrarily) slightly lower weight is dropped in the recursive feature elimination, the remaining feature will absorb its weight in the subsequent recursions, and could end up being labeled much more useful in the final rankings. Because the large n in this study means

there are plenty of observations to expend, resampling methods provide a simple route to improving the stability of these rankings. We chose to use the *multiple* SVM–RFE extension described by Duan *et al.* (2005) (not to be confused with multiple *class* SVM–RFE (Zhou and Tuck, 2007)), which imposes a resampling layer on each recursion pass such that the weights used for feature ranking/dropping are stabilized by averaging across multiple subsamples. Specifically, we used 10-fold cross validation as our subsampling method.

Exponential reduction in feature number: While the initial SVM–RFE paper described a one-by-one backward feature elimination, this can become time consuming when dealing with many features. Additionally, we often don't even care about the precise rankings among a bunch of useless features anyway. Therefore it can be useful to initially drop a greater number of the total features on each pass until the list of remaining features drops to a more manageable level. We chose to drop half of the features on each pass until the remaining number dropped below 5000. Then we switched to the one-by-one mode to give the most accurate rankings of the top features. While this can be detrimental in some situations (*e.g.* when a useful feature is dropped in a large batch, but its weight might have actually improved if some of those other features were dropped first) this short-cut was useful given restrictions on computing resources.

7.4.2 Optimal subset selection

Once we obtained ranked lists of the features for each imaging modality, ordered by their usefulness as judged by mSVM–RFE, the next step was to choose the optimal subsets of those top features. This is an important step in optimizing many types of machine learning classifiers, as you'd like to keep enough features to capture the most important aspects of the data (with respect to classification), but not too many as to lead to overfitting and poor generalization performance. Often there is some optimal middle ground.

At this stage, we shifted to a radial-basis-function SVM (RBF SVM) as our main classifier. This

is similar to the linear SVM used in the feature ranking algorithm, but employs a kernel transformation to allow for a nonlinear decision boundary in the original feature space (Compare Figure 7.6 to Figure 7.5). The width of the kernel parameter, together with the soft margin misclassification penalty, are optimized using grid search and internal 10-fold cross validation. This prevents overfitting the data. For example, if the decision boundary was allowed to be curvy enough, it would fit the training data perfectly, but would likely generalize poorly to the rest of the population.

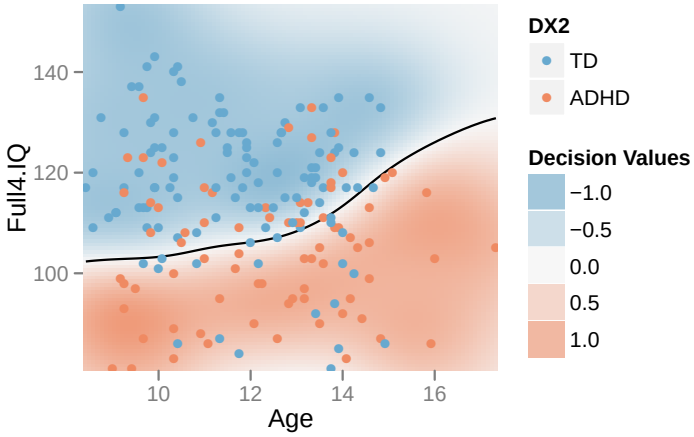
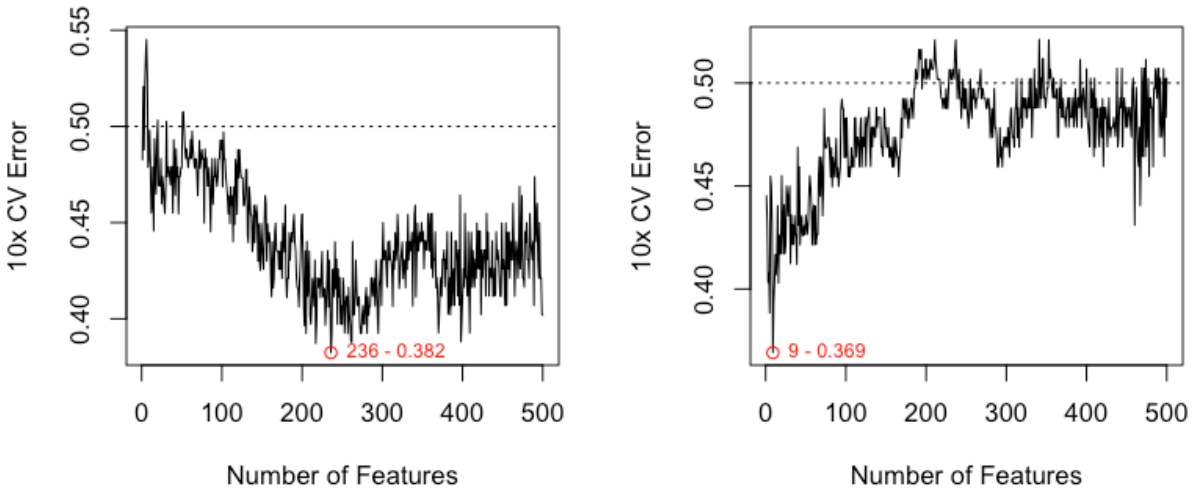


Figure 7.6: RBF SVM fit to Peking age and IQ data

Estimated generalization performance was determined using a layer of 10-fold cross validation. Within each fold, the classification accuracy on the hold-out samples was repeatedly gauged, while varying the number of the top features used as input. Averaging across all 10 folds allowed us to generate plots of generalization performance vs. number of features (Figure 7.7). The minimum along the curve was selected as the optimal subset of features to use in our final classifier. Importantly, this estimation of generalization performance and 10-fold cross validation was *external* to the feature ranking step. In other words, the features were ranked 10 times, each time independent of the hold-out samples for that given fold. This ensures that the estimated generalization performance is unbiased by spurious features that might nicely explain the training labels but don't generalize to the population (Ambroise and McLachlan, 2002). This is especially important when dealing with many features, since by random chance there will be some useless features that just happen to closely match the training labels.



(A) fMRI example

(B) Freesurfer example

Figure 7.7: Expected generalization error vs. number of top features

7.5 Final classifier

7.5.1 Site-specific classifiers

Our general approach to classification was to employ a collection of site-specific classifiers, rather than a single classifier with site as a feature. This allowed us to tailor our feature selection (§ 7.3) to each site, while also accommodating the unique site-specific aspects of the data set (§ 7.2.5). For example, because of varying T1-weighted scan qualities, the Freesurfer features might be very useful at one site, but not at another. Additionally, this approach is able to handle missing features (*e.g.* IQ is not reported from the NeuroIMAGE site), since the classifier for that site can simply be trained without them. For the Pittsburgh site, although there were no ADHD subjects in the training data, the available TD subjects were used to align TD feature means across sites, and thus allowed us to tap into the discriminating aspects of the data from the other sites (Figure 7.8). For the Brown site, which was most challenging because it lacked *any* training data, a similar across-site classifier was used, but it lacked the feature-wise bias adjustment that was made possible for the Pittsburgh

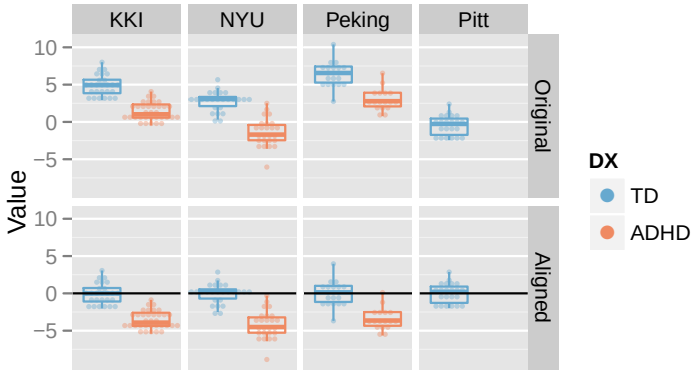


Figure 7.8: Across-site alignment schematic. For sites without complete training data (*e.g.* Pittsburgh), sites were aligned by their available subgroups (here, TD) and then an across-group classifier was trained.

site by its available TD training data. The Washington University site was excluded altogether from our final classification approach.

7.5.2 Combining modalities

Feature ranking (§ 7.4.1) and optimal subset selection (§ 7.4.2) were performed independently for each imaging modality. We chose this approach so that the tens of thousands of fMRI features would not swamp the much fewer Freesurfer features, and also so that we could choose the most effective processing options (*e.g.* 200 nodes vs. 400 nodes for extracting graph theory metrics from the fMRI time series data) among different preprocessing runs for the same modality. For each site, a set of RBF SVMs were then trained — once for each imaging modality, once for the demographics alone, and once with all the top features from all modalities together. These were used to generate a list of class predictions for each test set subject. Final assignment was made by simple majority voting. ADHD subtypes were assigned according to what was most common in a given site's training data (Tables 7.1 and 7.3).

7.6 Discussion

7.6.1 Feature selection

The robustness of our feature selection approach is one positive aspect to take away from our effort. The multiple SVM-RFE method was able to handle large feature sets (*e.g.* the CC400 atlas has 160,000) in a reasonable amount of time, and consistently returned useful ranked lists of the top features. As a reality check, consider the IQ feature, which, from simple inspection (Figure 7.1), we already knew was a *highly* useful feature. Therefore we expected it to rise towards the top of our ranked feature lists when it was included with features from the other imaging modalities. However, a worry was that maybe highly useful features like IQ could get lost among the sea of useless features that we expected to be in the raw output of the fMRI processing. When the feature ranking was performed, however, IQ consistently rose into the top 10 list of features — often *the* top feature — even when accompanied by *all* 160,000 CC400 features and no prior constraints. On the flip side, this early observation was already a clear sign to us that the imaging features were not going to be as marginally useful as we all had hoped.

7.6.2 Creative features

The motion features we generated are another clear example of the fact that just because a feature shows significant differences between classes does *not* necessarily imply that it is going to improve classification performance. This is especially important to remember for studies with high n, since it is easy for even small effect sizes to be "highly significant". Even though these are true effects, and perhaps interesting to study in population-based research studies, their small size and large individual variability combine to limit their usefulness for the diagnostic classification of individual subjects.

7.6.3 Collaboration and data sharing

The collaborative data sharing experience of the ADHD–200 initiative is one of its most important and widely-applicable outcomes. On one level, the collaboration between the organizing sites, together with their open-access philosophy toward data sharing, has combined to produce one of the most valuable publicly-available developmental neuroimaging data sets. Even more exciting is the response that the community has had in building on this foundation. We of course make our tools available to anyone who wants to use them (`http://github.com/johncolby/SVM-RFE`), but the best example is the preprocessed version of the original fMRI data set that has been contributed back to the community by the Neuro Bureau (`http://neurobureau.projects.nitrc.org/ADHD200`). This resource lowers the entry barrier to the field, allowing a broader range groups (*i.e.* those lacking an fMRI background) to start exploring the data without having to get bogged down in the details of fMRI preprocessing. Additionally, for those already in the neuroimaging field, it brings a sorely-needed level of standardization between studies, which allows for an enhanced ability to replicate results (or compare differences) between different analyses. After all, there are so many different processing parameters in modern neuroimaging studies, that, even if two groups started from the same raw data, there will always be the nagging suspicion that different findings could simply be due to different processing settings. With this type of standardized preprocessed data made available, however, we can be more confident that any observed differences are due to the specific parameter we are actually trying to investigate — in this case our classification methodologies.

7.6.4 Multisite studies

Another important and generalizable area of insight concerns the design of multisite studies. On their face, they sound like a great idea: More collaboration, higher n, more power, broader generalization, and distributed costs and responsibility. However, all of these positive gains must be weighted against the negative impact of increased variability in diverse aspects of the data due to site-specific effects. One striking example from the ADHD–200 data set is that the distributions

of one of the broadest anatomical measures — mean cortical thickness across the entire brain — varied widely between sites. Knowing that ADHD is associated with thinner cortex (Shaw *et al.*, 2006b), you might naturally train a classifier to diagnose based on this feature. However, if you applied this classifier to subjects from a different site (one that has a global negative site effect on thickness) without accounting for this overall site offset, you might inadvertently misclassify all of those subjects as having ADHD. This situation is demonstrated using simulated data in the schematic of our between-site feature alignment procedure for the Pittsburgh site (Figure 7.8). If the Pittsburgh TD subjects were classified based on a classifier trained on this simulated feature at the KKI site, they would all be erroneously classified as ADHD due to a simple negative site offset at Pittsburgh. Another example is the IQ score, which was derived from different raters and different test instruments depending on site. Further, just like in this project, different sites could be missing different aspects of the data. In the end, we opted for a series of site-specific classifiers rather than a single across-site classifier. In effect this largely forgoes the potential benefits of a true multisite study, and instead could be better described as a meta analysis of ADHD classification based on a handful of individual studies. The point, then, is that it is crucial to strive for across-site uniformity in multisite studies, as it is all to easy to let a handful of uncontrolled parameters spoil their potential benefits.

7.6.5 Performance on test data

The most surprising result to come out of the ADHD–200 competition was that, although imaging features were moderately useful for classifying ADHD from TD subjects, including these features failed to provide any additional benefit over using demographic features alone. There were 195 possible points (1 test subject was excluded), which would have required correctly predicting all TD subjects as well as all ADHD subtypes. The winning *imaging*-based classifier scored 119 points. For reference, we scored 110.5 points, which put us in 4th place out of 21 entries. However, the best *overall* score of 124 points was reached by completely ignoring all of the imaging features and relying solely on the demographics information. Even though this is close to the

195 points $\times$ 65.5 % = 127.725 points we predicted based on our own analysis of the demographics features in § 7.2.6, we and others decided to stick it out with the imaging features since that was the focus of the project. Still, this finding makes the cautionary point that when we see claims like "Feature X is useful for classifying disease Y", we must always ask the question, "Relative to what baseline?" Similarly, it is also important to consider the performance we could have expected based on chance alone. Based on hierarchically flipping a coin to determine TD vs. ADHD, and, if the diagnosis is ADHD, flipping it again to determine the ADHD subtype diagnosis, the study organizers describe an expected chance-level accuracy of 38.75 % (corresponding to 86.5 points). However, as we've discussed before, a more useful baseline is the no-information rate, which takes into consideration the actual class prevalences in the training set. Seeing as TD was the most common overall diagnosis in the training set, if we had simply predicted *all* of the test subjects would also be TD, we would have received 195 points $\times$ 55 % = 107.725 points. With respect to the performance of the imaging-based classifiers, this number is just as striking as the demographics-only performance. It suggests that around half of the imaging-based entries would have performed better — and certainly expended less effort — by simply classifying everyone as TD. Considering the large individual variability that accompanies even real group effects, and the relatively low prevalences that exist in community populations (compared to research studies), it is unlikely that neuroimaging features are going to be useful any time soon for the *diagnostic* classification of behaviorally-diagnosed neurological syndromes like ADHD — and, in the case of ADHD, why would we even want them to? Existing diagnostic instruments are quick, effective, and cheap (*DSM-IV* 2000; Goldman *et al.*, 1998; Power *et al.*, 2001), and existing treatments are effective (Elia *et al.*, 1991) and with relatively low risk (Cooper *et al.*, 2011; Habel *et al.*, 2011). Rather, imaging-based classifiers of ADHD show their biggest promise in populations where the pretest probability is high or diagnosis is already assumed. For example, in a scientific context towards further understanding the neurobiological basis of the disorder, these techniques can be used to map regions of the brain that are most useful for classification, thereby providing a complementary tool to standard hypothesis testing. Similarly, in a clinical context, they may be more useful in predict-

ing diagnostic subtypes among individuals who have already been screened from the population at large, or for predicting later treatment response and prognosis.

7.6.6 Conclusion

On the whole, the ADHD–200 project was able to successfully demonstrate that imaging features show modest potential for the classification of ADHD subjects from typically-developing controls. Additionally, the initiative provides a strong example of the benefits of collaborative open-access research and data sharing within the neuroimaging community. However, the biggest message is that we must be realistic in our expectations about what neuroimaging data can — and cannot — do for us. This way our limited effort and funding can be best targeted to the most high-yield types of research questions.

CHAPTER 8

Comprehensive mapping of white matter development in the PING sample ($n = 869$)

8.1 Introduction

Our previous discussion has shown that the trajectories of many aspects of human brain development are nonlinear in nature, with developmental changes occurring rapidly from before birth, and then slowly decelerating with age. Further, we have seen how these phenomena vary by brain region, and exhibit regionally-specific relationships with cognitive measures (Chapter 1). Continuing our focus on the white matter, it is also clear that different white matter tracts develop with different timings in their developmental trajectories (Figure 1.7; Lebel *et al.*, 2008b), and that white matter development generally takes place in a caudal-to-rostral gradient across different brain regions (Chapter 4). Knowing that white matter properties also vary prominently *along* specific white matter tracts (Chapter 5), we set out to provide the first along-tract maps of white matter developmental timing (Aim 1b).

While we had originally aimed to carry out this exploration on our UCLA data set, which would have allowed us to also study how these timing patterns are possibly affected in the context of FASDs, we were instead given the exciting opportunity to conduct this investigation on the Pediatric Imaging, Neurocognition, and Genetics (PING) data set. Similar to the ADHD–200 effort, PING is a large, multisite, and multimodal study that, among its aims, will also provide a valuable resource to the community through public data sharing. The broad goal of PING is to map the genetic, environmental, and neurobiological factors that contribute to our "individuality". These

aspects are being assessed in a large cohort of 1400 individuals (age 3–21 years) via a genome-wide single nucleotide polymorphism (SNP) array (Affymetrix 6.0; Bakken *et al.*, 2012), a background questionnaire on demographics and medical history, brain imaging (including a T1-weighted high resolution anatomical scan, a resting-state functional MRI scan, and a 30-direction DTI scan), and a computerized neuropsychological and cognitive testing battery. This initial effort will focus on mapping patterns of variability in the developmental timing of white matter maturation, but this rich data environment will also allow us to eventually generate novel along-tract maps of relationships with the genetic and behavioral data.

8.2 Demographics

DTI data have been collected for 869 participants as of January 2012. Out of these, 855 are "complete" — indicating that they have also completed all the other aspects of the study. This subset of subjects is being used for further analysis. 9 of the 10 PING sites are represented in this sample, including Cornell University, UC Davis, University of Hawaii, the Kennedy-Kreiger Institute at Johns Hopkins University (KKI), Massachusetts General Hospital at Harvard University (MGH), UC Los Angeles (UCLA), UC San Diego (UCSD), University of Massachusetts Medical School (UMMS), and Yale University. Data were collected between January 2010, and August 2011, with different sites joining the project at different times (Figure 8.1).

8.2.1 Gender

Genders are represented relatively equally, and the ratio between males and females does not significantly vary by site ($\chi^2_8 = 2.04$, $P = 0.98$) (Table 8.1).

8.2.2 Age

Subjects are between the ages of 3 and 21. Males (range 3–21, mean 13.1 ± 4.8 years) and females (range 3–21, mean 13.3 ± 5.0 years) do not significantly differ in age ($F_{1,836} = 0.25$, $P = 0.62$).

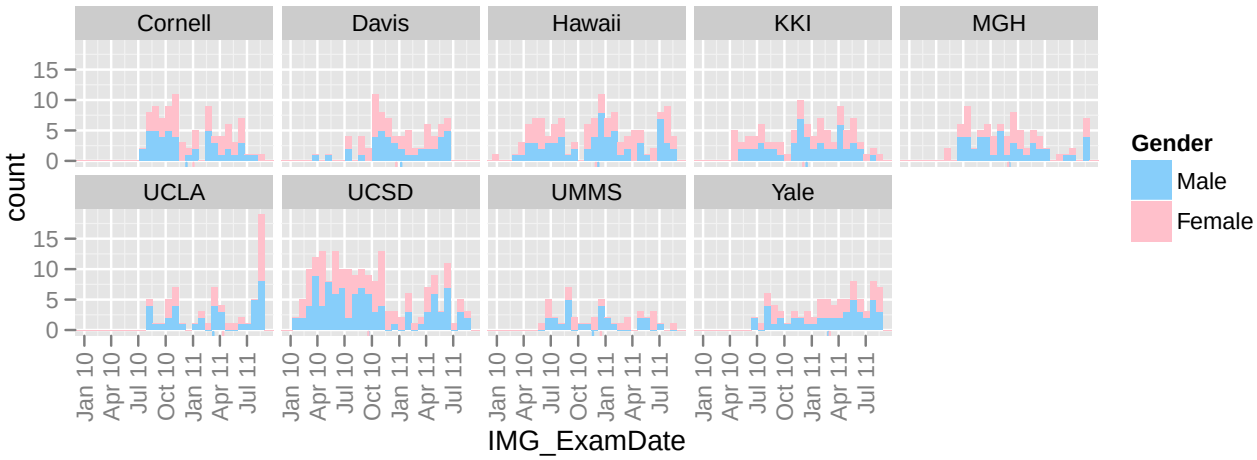


Figure 8.1: Distribution of subjects by recruitment date (`IMG_ExamDate`) and gender. Gender means are shown as tick marks in the plot rug region.

	Gender		
Site	Male	Female	**Sum**
Cornell	45	47	92
Davis	40	38	78
Hawaii	65	61	126
KKI	51	49	100
MGH	44	37	81
UCLA	38	27	65
UCSD	99	89	188
UMMS	24	19	43
Yale	45	37	82
Sum	451	404	855

Table 8.1: Number of subjects by gender and site

However, there is clearly a site effect on age ($F_{8,836} = 7.88$, $P = 3.02 \times 10^{-10}$). For example, the mean age at the KKI site (15.4 years) is over 3 years older than at the UCSD site (12.1 years). Additionally, while there is no main effect of gender on age, there is a significant site:gender interactive effect ($F_{8,836} = 2.03$, $P = 0.04$). For example, on average, females are 2.5 years *older* than males at UC Davis, but females are 1.5 years *younger* than males at MGH.

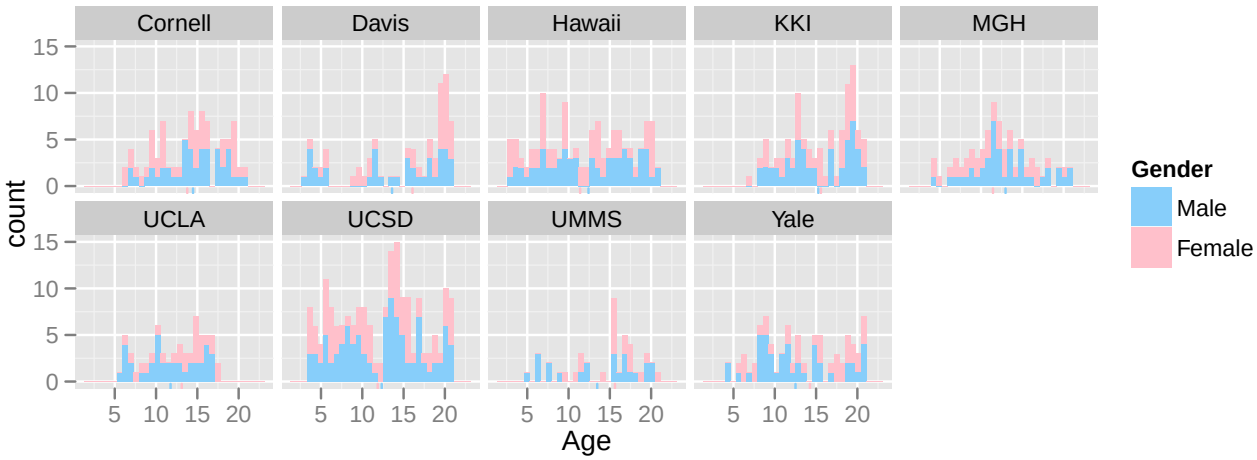


Figure 8.2: Distribution of subjects by age and gender. Gender means are shown as tick marks in the plot rug region.

8.3 Methods

Different sites have different MRI scanners, so there was variation in the exact DTI sequence used. However, more general parameters were aligned across sites, including the use of 30 diffusion-encoding directions and $2.5 \times 2.5 \times 2.5$ mm isotropic voxel size.

8.3.1 Preprocessing

Several initial preprocessing steps were performed on the DTI data as part of a broader study-wide preprocessing pipeline at UCSD. These included 1) conversion of the scanner-specific raw data into a common NIfTI data format, 2) mapping and unwarping of eddy-current-induced geometric distortions in the images through the use of multiple b0s with opposite phase-encoding directions, 3) registration of the diffusion-encoding volumes to the $b = 0$ volume, 4) registration of all subjects' DTI data to a common template space, and 5) calculation of a binarized brain mask.

8.3.2 Automated atlas-based deterministic tractography

Once the data were transferred to UCLA, we implemented a fully-automated atlas-based deterministic tractography protocol to extract tract groups for all subjects. This approach contrasts with the

manual method used in Chapters 5 and 6. For each subject, a standard DTI model of diffusion was fit to the raw data, and then whole-brain deterministic tractography was performed. General tracking constraints included an FA threshold of 0.15, a fiber turning angle threshold of 60°, and a mask to exclude fibers outside the brain. These steps were performed with Diffusion Toolkit v0.6.2. Next, FLIRT from the FSL toolkit was used to register the UCSD template image into MNI152 space using a 12-parameter affine approach. This transformation matrix was then applied to each subject's fiber data. According to the Wakana *et al.* (2007) protocol, a single set of tracking ROIs was drawn for the left hemisphere tracts. The FMRIB58 FA template in MNI152 space was used as the reference image. These ROIs were reflected across the mid-sagittal slice to generate an additional, unbiased, set of ROIs for the right hemisphere tracts. Since the streamline data for all subjects were already in the same space, these ROIs could then be automatically applied to mask out all of the tracts of interest. This was performed with TrackVis v0.5.2. Additionally, for each subject, we automatically compiled and annotated these ROIs and tract groups into a TrackVis "scene" for later interactive viewing (Figure 8.3). For example, the arcuate fasciculus and its tracking ROIs are shown in red in the 3D display, as well as labeled with their names and colored red in the list of annotations to the right.

8.3.3 Extraction of *along*-tract data

As with the tractography processing, the common template space also facilitates automating the along-tract data extraction step. For each tract, a tract origin was specified in template space (*e.g.* a point in the frontal lobe was chosen for the left inferior frontal occipital fasciculus; See Figure 5 in Colby *et al.* (2012) for the rest of our conventions). The streamlines in that tract group for *all* the subjects were then reoriented accordingly. Spline-based resampling of the underlying scalar volume, collection of data across streamlines, calculation of the mean streamline geometry, and final data export then proceed as described in § 3.3.7 and Chapter 5. For each subject, the mean tract geometries (*i.e.* the tract "cores") and along-tract scalar estimates were automatically compiled into a custom TrackVis scene for later visualization (Figure 8.4).

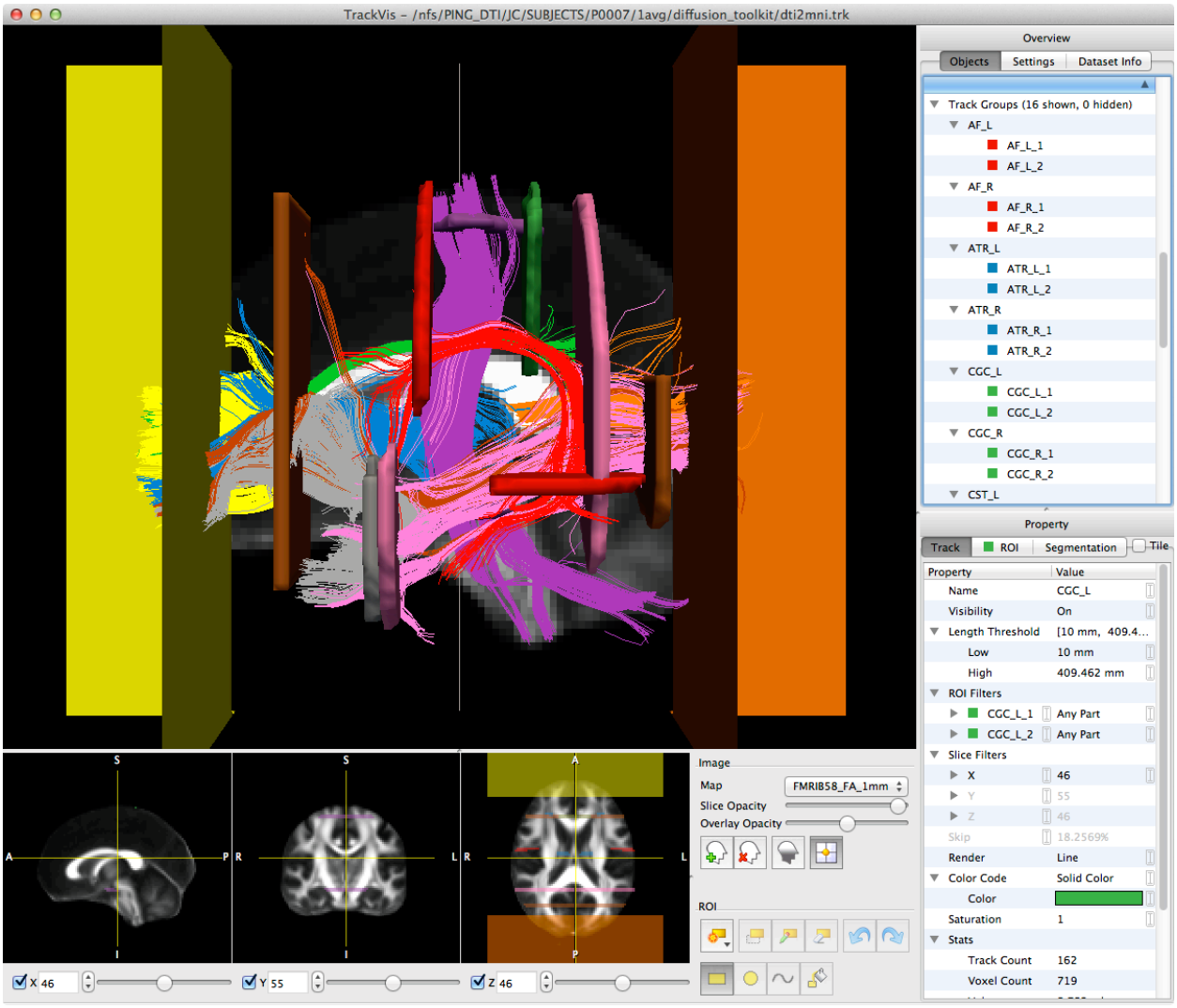


Figure 8.3: Example automatically-generated TrackVis scene

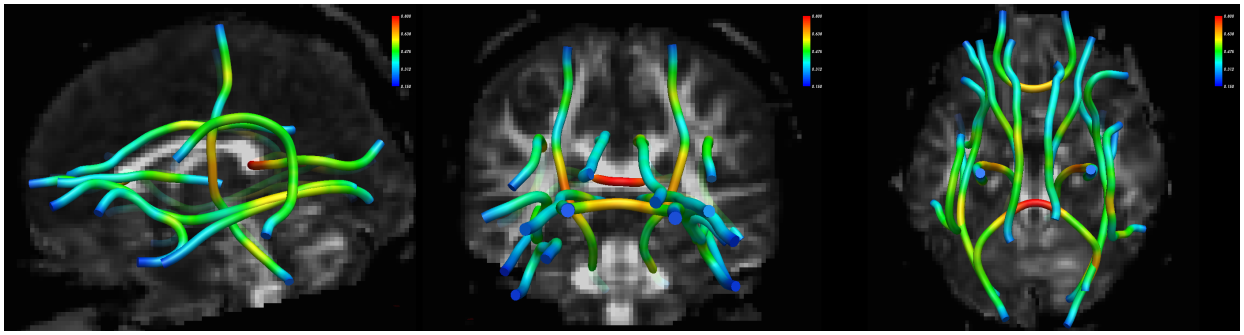


Figure 8.4: 3D rendering of the along-tract variation in FA for a single subject. The scalar estimates are overlaid onto the mean tract geometries for this subject. Warmer colors represent higher FA.

8.3.4 Quality control

It is important that any neuroimaging processing pipeline be as automated as possible. This minimizes human errors, and is often absolutely necessary from a practical standpoint. Such is the case with the PING data, which include 869 subjects and potentially 16 tract groups each. Accordingly, we designed the previous steps to be fully-automated and embarrassingly-parallel across subjects. Total processing time per subject is approximately 3 minutes, and all subjects can be submitted simultaneously to a cluster job scheduler. This means that the final processed data, ready for statistical modeling, can be regenerated from scratch in less than an hour.

Nevertheless, it is still crucial that quality control measures be put in place to ensure the validity of the processed data, which will eventually be analyzed, interpreted, and reported to the community. We implemented several QC steps into our methods:

Tracking failures: As described in § 6.3, the number of tracking failures is a simple — yet important — metric to examine. Since PING is a multisite study, we wanted to compare the tracking success rates between sites in order to determine if there were any site-specific processing issues. We did this by tabulating, for each site, the distribution of the number of tracking failures per subject. There were 16 tract groups in this analysis, so 16 failures would mean the automated tract delineation protocol didn't identify a single tract. Such a situation in an individual subject would make us look closer at their data for problems, and a pattern of many subjects with many failures at a given site would make us look closer at that site for

problems. Early on in our analysis, this approach let us identify a significant problem with 6 out of the 9 sites (Figure 8.5). It turned out that these sites all had Siemens Trio scanners, and the problem was traced to an error in handling the diffusion-encoding gradient table from this type of scanner in the early preprocessing pipeline at UCSD.

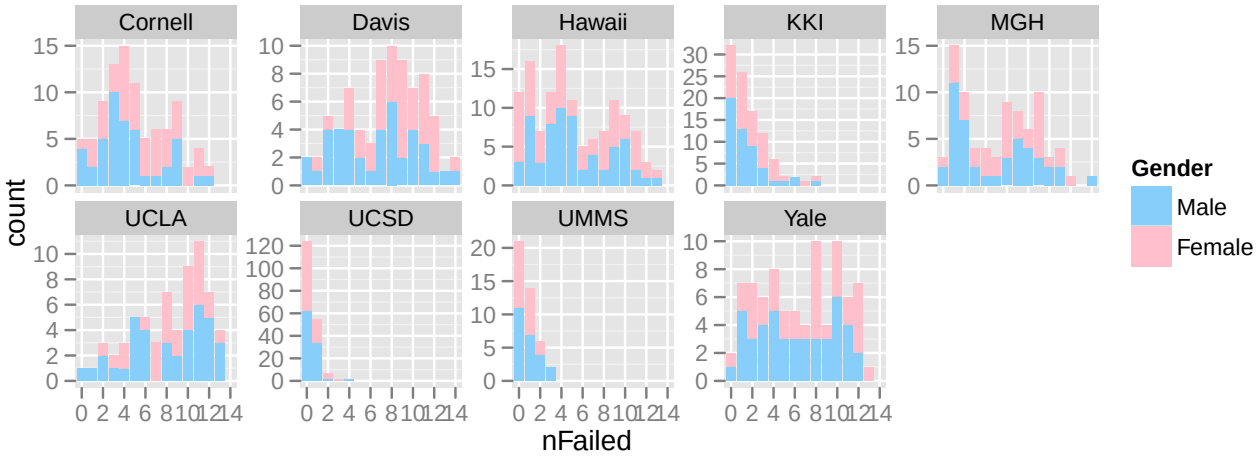


Figure 8.5: Distributions of tracking failures (lower `nFailed` numbers are better) by site and gender

Streamlines: In order to optimize our automated tracking protocol, we also examined, for each tract and hemisphere, the distribution of the number of streamlines in the tract dissections (Figure 8.6). This makes it easy to identify if there are any major problems (*e.g.* incorrect tracking ROIs or incorrect filters on length).

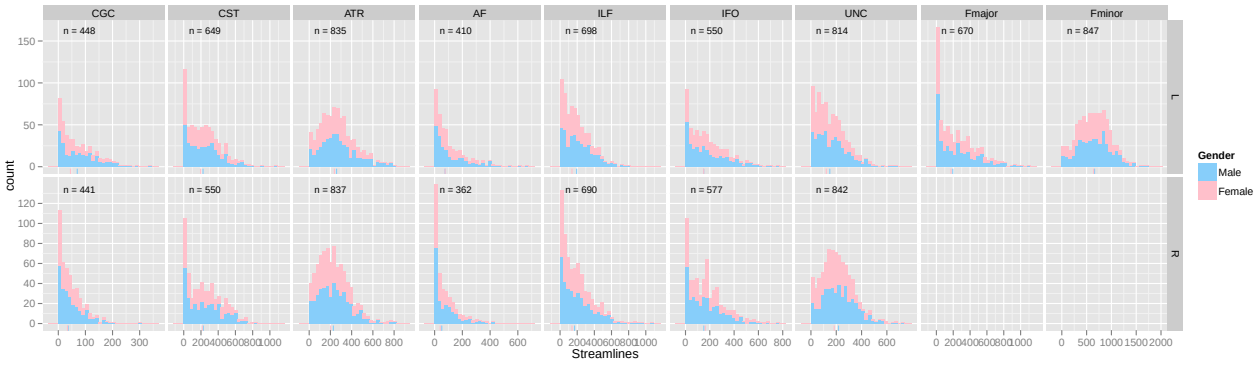


Figure 8.6: Distributions of the number of streamlines, by tract, hemisphere, and gender

Raw along-tract data: Similarly, examining the raw along-tract data is a good way to see if there are any questionable sub-populations for a given tract group. In this analysis, the *forceps*

minor dissections initially contained an obvious sub-population of streamlines with lower FA. Closer examination traced the source of the problem to cerebellar fibers that were being grouped in, and allowed us to then tweak our tracking protocol to address the issue.

Correspondence plots: The usefulness of this analysis also rests largely on the accurate extraction of along-tract data for each of the tract groups generated with the automated tractography protocol. As we have discussed previously, correspondence plots are an efficient way to supervise this process, as they show which streamlines will be grouped together along the entire tract (*e.g.* Figure 9 in Colby *et al.* (2012)). For each subject, a HTML QC report was automatically generated containing an overview of all of the correspondence plots for a given subject (Figure 8.7). Additionally, a more detailed view automatically appears when you hover over a specific tract (Figure 8.8).

Video loops: With 869 subjects, even glancing at a simple QC HTML page for each subject takes a lot of time. Therefore, the last, and most effective, QC measure we employed was to view all subjects' QC images for a given tract in a video loop at 15 frames per second. This allowed us to directly see the data for all individual subjects in about one minute. Surprisingly, even at the speed, the brain is very effective at identifying outlying individuals. These subjects can then be investigated closer to see if any changes need to be made. As an example, we include a QC video for the left arcuate fasciculus (`http://vimeo.com/johncolby/qcvideoloop`), which includes examples of both tractography and along-tract data extraction errors, as well as several suspect raw images.

8.3.5 Statistical modeling

Across-tract properties (tract length, tract-averaged FA, number of streamlines) were examined first, using mixed-effects linear models and ANOVA to examine main effects of tract, hemisphere, gender. Typically, a single subject-level intercept term was included as a random effect to accommodate the repeated measures in the data. Interactive effects between predictors were also

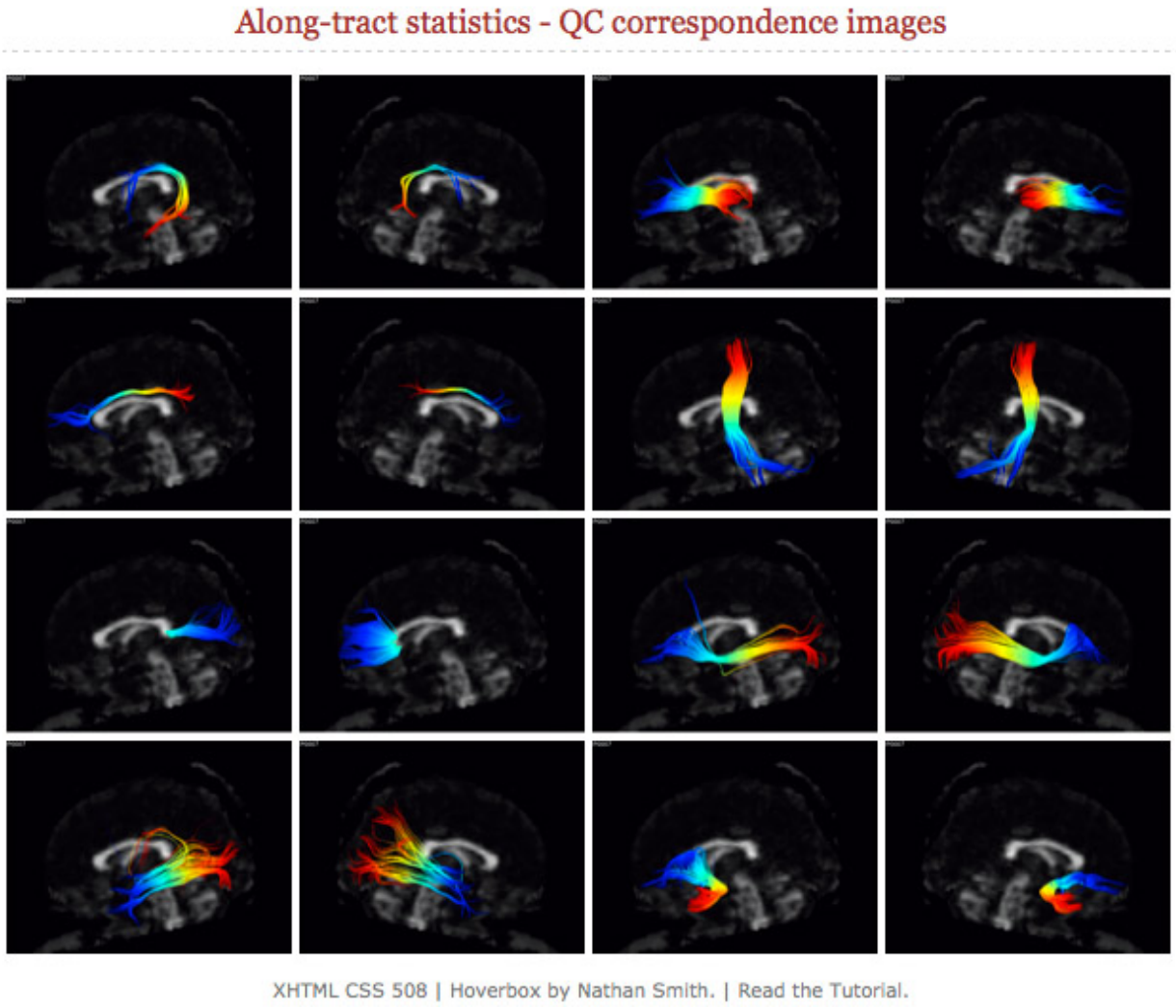


Figure 8.7: Automatically-generated HTML QC report for along-tract data

considered.

For the along-tract data, a mono-exponential model of development was fit to the FA vs. age developmental trajectory for all subjects' data using nonlinear least squares. This was done separately for each combination of tract, hemisphere, gender, and along-tract position. This model assumes some initial FA value at birth (FA_0), and then approximates white matter development — in terms of increasing FA — as decelerating with age until an eventual adult plateau is reached (FA_∞) (Figure 8.9). The third parameter in this model is a unitless exponential decay constant (λ), which describes whether these developmental changes occur rapidly or slowly. The decay constant is often reformatted into a time constant, for example $t_{75\%} = \frac{\ln(4)}{\lambda}$, which is easily interpretable as

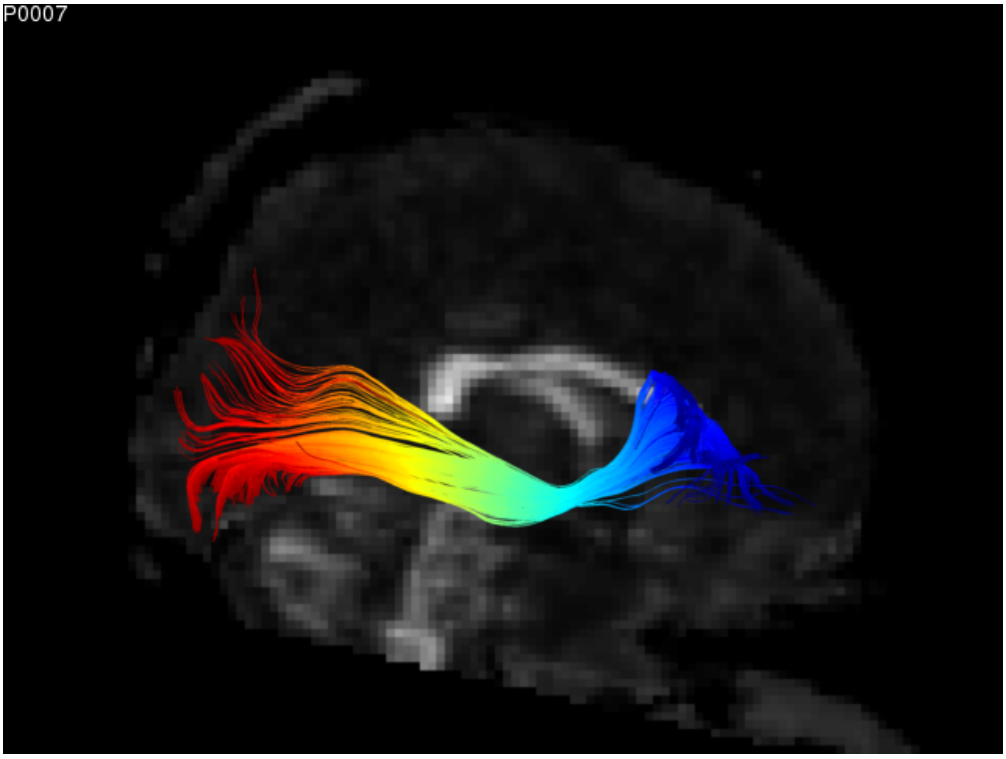


Figure 8.8: Example correspondence plot (R IFO). Color hue designates which vertices will be grouped together for the analysis.

"the time it takes (in years) for 75% of the developmental changes that will ever occur from birth to plateau, to have occurred".

To stabilize the parameter estimates in these nonlinear fits, the process was wrapped in a bootstrap resampling scheme and replicated 1000 times.

8.4 Results

8.4.1 Whole-tract observations

In MNI152 space, the average tract lengths ranged from 60 mm for the uncinate fasciculus to 151 mm for the inferior frontal occipital fasciculus. Like in Chapters 5 and 6, these lengths were used to adaptively set the number of interpolation points for extracting along-tract data (Table 8.2).

On average, compared to the left hemisphere, tract groups in the right hemisphere were not

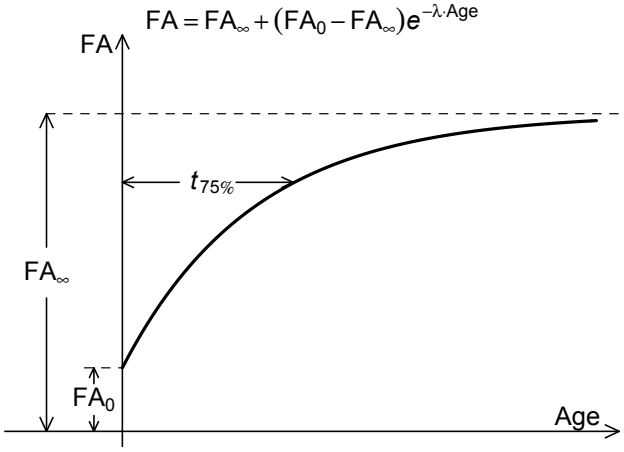
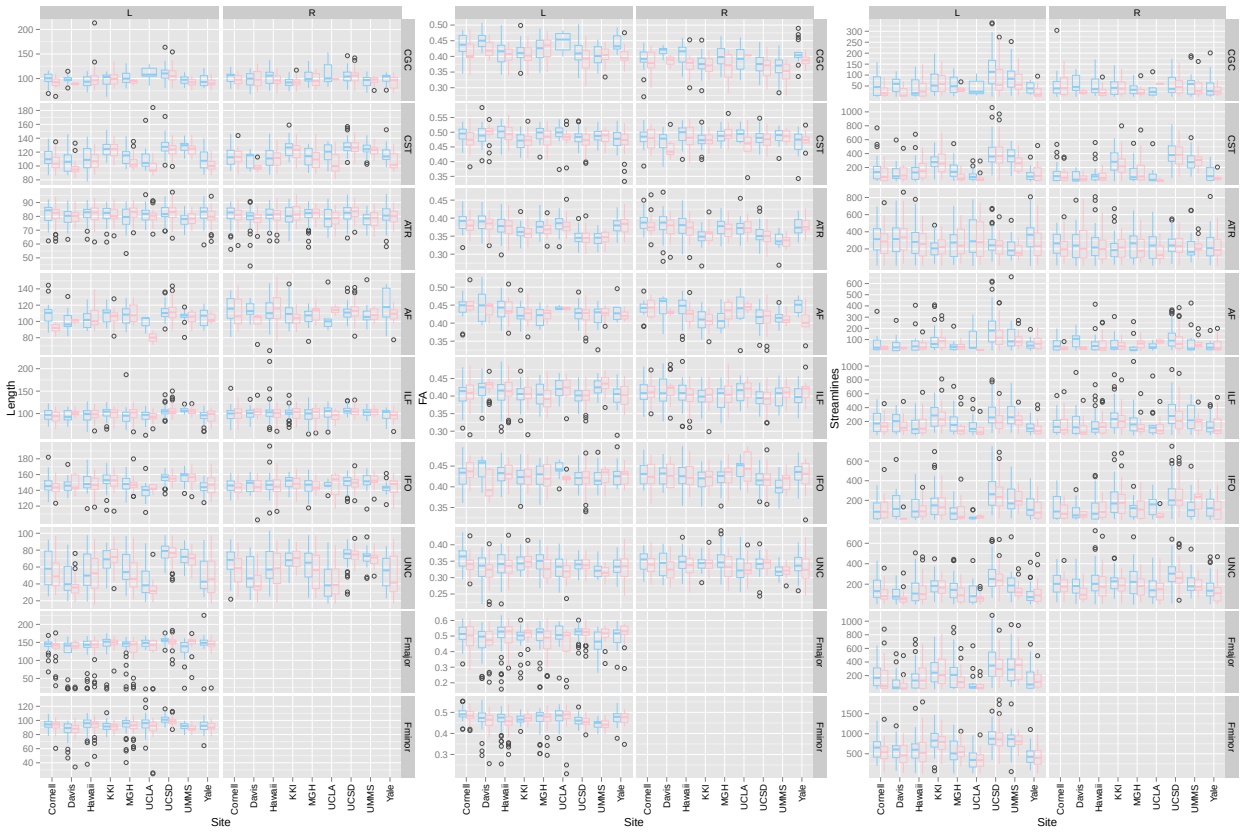


Figure 8.9: Mono-exponential model of developmental increases in FA

significantly different in length ($t_{7745} = 0.35, P = 0.73$), and did not have a significantly different number of streamlines ($t_{7745} = -1.15, P = 0.25$). However, they did have 0.008 lower FA ($t_{7745} = -7.27, P = 3.9 \times 10^{-13}$). On average, compared to males, tract groups from females were 2.5 mm shorter ($t_{853} = -2.64, P = 0.008$), had 0.005 lower tract-averaged FA ($t_{853} = -2.93, P = 0.004$), and had 20 fewer streamlines ($t_{853} = -3.73, P = 2.0 \times 10^{-4}$). The distributions of these metrics, grouped by tract, hemisphere, and gender, are displayed in Figure 8.10.

8.4.2 *Along*-tract variations in developmental timing

We observed wide variation in developmental timing when the $t_{75\%}$ parameter was qualitatively examined between tracts, hemispheres, along-tract positions, and genders. Using our previously-developed visualization tools for along-tract analyses, we generated 2D (Figure 8.11) and 3D (Figure 8.12) plots of these results. When averaged across all the data, the average $t_{75\%}$ was 9.2 years. Developmental timing was 1.7 years earlier in the right hemisphere tracts than their left hemisphere couterparts ($t_{1085} = -4.07, P = 5.09 \times 10^{-5}$). Females had earlier developmental timing than males ($t_{1085} = -9.60, P = 5.47 \times 10^{-21}$), and this effect varied by hemisphere ($t_{1085} = 4.30, P = 1.90 \times 10^{-5}$). In the left hemisphere, females had developmental timing 6.4 years earlier than males, but in the right hemisphere the difference was only 2.3 years earlier.



(A) Length **(B)** FA **(C)** Streamlines

Figure 8.10: Distributions of whole-tract properties, by tract, hemisphere, and gender

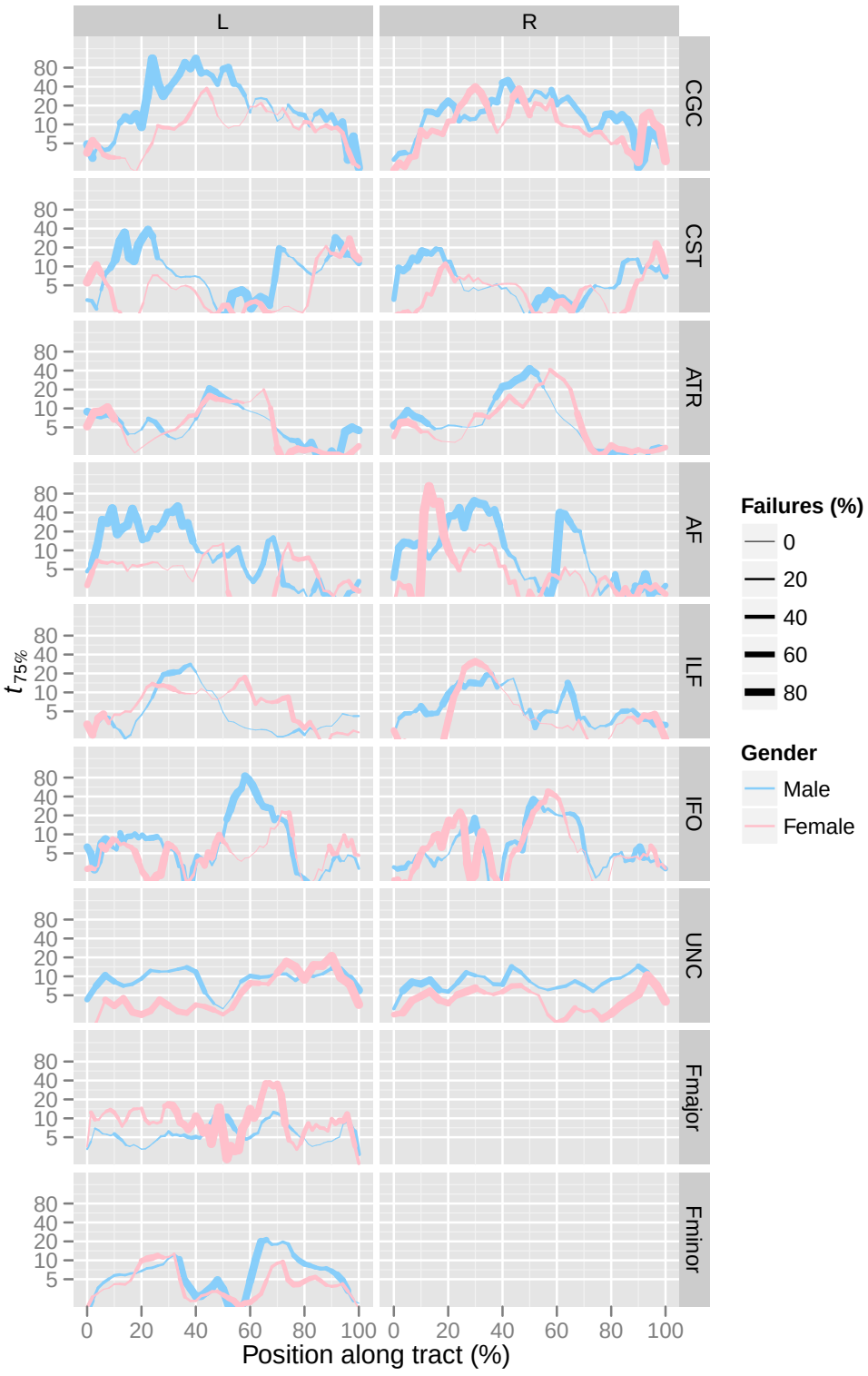
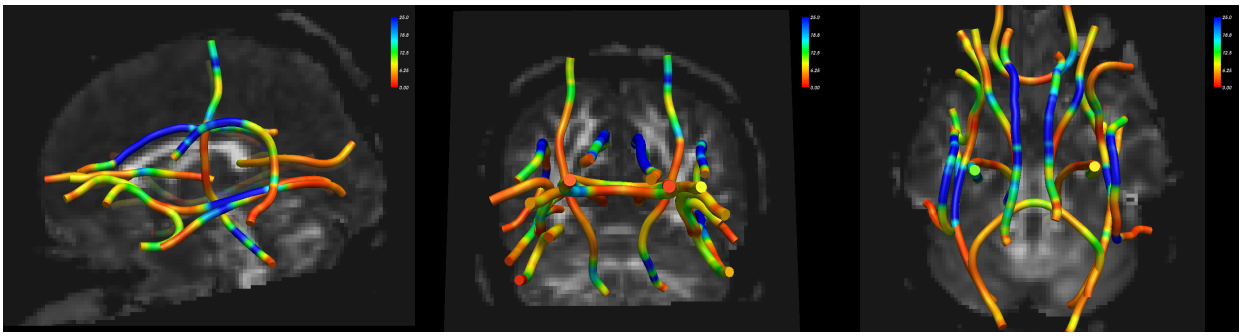
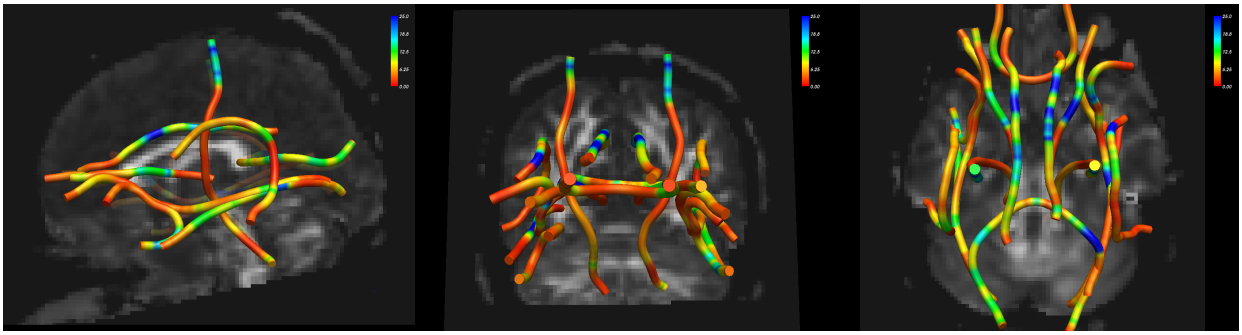


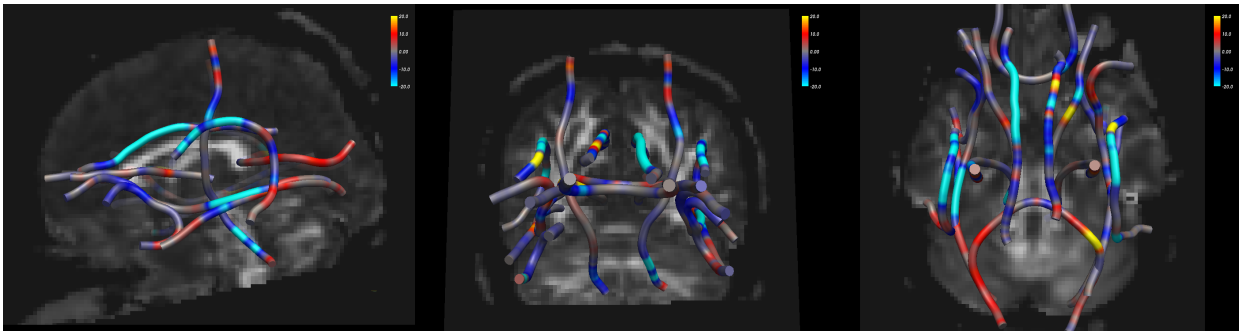
Figure 8.11: Along-tract variation in developmental timing (2D), by tract, hemisphere, and gender. The y-axis is on a $\log_2$ scale and is in units of years.



(A) $t_{75\%}$ in Males



(B) $t_{75\%}$ in Females



(C) Female – Male difference map

Figure 8.12: Along-tract variation in developmental timing (3D), by tract, hemisphere, and gender. In (A) and (B), cooler colors represent slower developmental timing. In (C), cooler colors represent slower developmental timing in males compared to females.

		Hemisphere	
	Tract	L	R
1	CGC	102.24 (51)	100.41 (51)
2	CST	116.42 (59)	120.03 (59)
3	ATR	81.22 (41)	80.53 (41)
4	AF	108.25 (55)	110.64 (55)
5	ILF	99.48 (51)	101.51 (51)
6	IFO	150.96 (75)	147.87 (75)
7	UNC	59.95 (31)	61.63 (31)
8	Fmajor	139.96 (71)	
9	Fminor	93.30 (47)	

Table 8.2: Mean tract lengths (mm) across all PING subjects. The numbers of interpolation points used for along-tract modeling are given in parentheses.

8.5 Discussion

8.5.1 Developmental timing

The finding of earlier developmental timing in females is consistent with previous reports on the timing of maturational changes in the white matter (Asato *et al.*, 2010; Herting *et al.*, 2011; Lenroot *et al.*, 2007), cortex (Clayden *et al.*, 2011; Giedd *et al.*, 1999a), and other brain structures (Giedd *et al.*, 1997). It is also in line with the much broader cross-sectional literature documenting gender effects on brain structure (Lenroot and Giedd, 2010). However, the really novel aspect of the present analysis is that we break this down further, and go on to examine how developmental timing — as well gender and hemisphere effects on this trajectory — vary along major white matter tracts in the human brain (Figures 8.11 and 8.12). When examining these plots, it is clear that some tracts show prominent along-tract variation in their developmental timing. To get a feel for what different $t_{75\%}$ values look like, in terms of their corresponding FA vs. age developmental trajectories, we plotted two different positions from the left inferior longitudinal fasciculus (ILF) in Figure 8.13. The more anterior position (32%) is relatively slowly-developing, and has $t_{75\%} = 25$ years, while the more posterior position (58%) is relatively quickly-developing, and has $t_{75\%} = 3$ years. The arcuate fasciculus is another tract that shows along-tract variations, with slower development in

its frontal lobe portion. This is consistent with the literature describing frontal lobe development to be especially protracted (Chapter 1). The anterior thalamic radiations (ATR) also show a more protracted time course as you move towards the frontal lobe from their thalamus terminus. This tract, in particular, is striking because these changes form a smooth gradient that is nearly monotonic through the posterior half of the tract. The ATR is also a useful example to show that some features of the along-tract pattern of developmental timing can be broadly conserved between hemispheres and genders. The along-tract variation in the gender effect is also prominent. For example, females show earlier timing at nearly every point along the left anterior portion of the cingulate (CGC), but not in the posterior half of the tract, and not in the right hemisphere. These differences are highlighted in Figure 8.12C. Along the corticospinal tract (CST) in both hemispheres, females also show earlier timing towards either tract terminus, but not in the central portion of the tract. Although still preliminary, these findings already suggest that there is wide variability in the FA vs. age developmental trajectory even within the same white matter tract in the human brain.

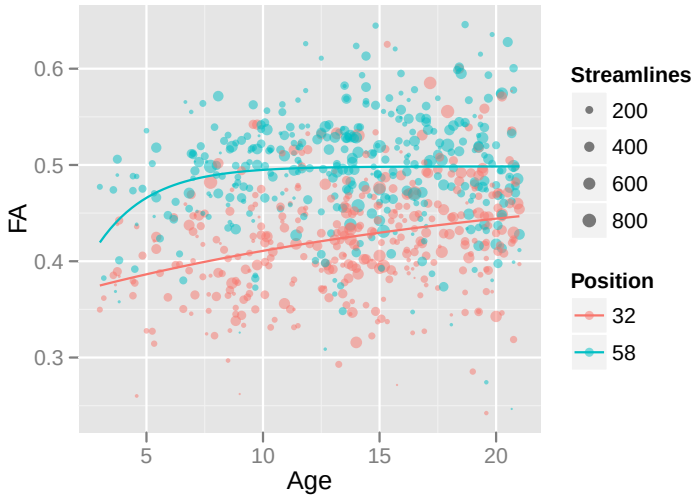


Figure 8.13: Example along-tract variation in developmental trajectories. FA is plotted vs. age for all subjects, for two different positions along the L ILF. These positions were chosen approximately one third (32%) and two thirds (58%) from the anterior tract terminus. Model fits for the mono-exponential developmental trajectory model are overlaid. The number of streamlines in each tract group is encoded by point size.

8.5.2 Exponential model

The choice of which developmental model to use is an important one. We decided to use a mono-exponential function because of its previous use in the developmental neuroimaging community (Lebel *et al.*, 2008b), and the easy biological interpretability of its three parameters. While this is often a useful simplification, it does ignore some of the nuances of brain development. As one example, this model assumes monotonically increasing FA towards an asymptotic plateau. In reality, many white matter regions actually exhibit modest decreases in FA during adulthood, which have been approximated by quadratic fits if the maximum age is younger (Colby *et al.*, 2011; Lebel and Beaulieu, 2011), or poisson fits if the maximum age is older (Lebel *et al.*, 2012). Additionally, unlike the simpler physical systems where these models originated (Michaelis *et al.*, 2011), there is no strong theoretical basis that human brain development should follow such simple kinetics. Our application of the same model to all brain regions is a similar trade-off. One the one hand, it certainly simplifies the analysis and interpretability of the results. Again, however, the true neurobiology is likely less homogeneous region-to-region. Methodological concerns come into play as well. From examining Figure 8.11, we can see many regions have a thick line, which indicates that there were many failures of the model fits across the subsampled data replications at these points. For some regions, this likely implies that the mono-exponential model is a poor approximation to the true developmental trajectory. For others, it means that our age range failed to capture the developmental changes and that the FA vs. age curve appears flat. In these regions, some of the nonlinear least squares fits could fail to converge, and, for the times when the model does converge, the estimate for the timing parameter will swing wildly from very early to very late. Similarly, we should be suspect of extremely small and extremely large values for the timing parameter that are far outside of our age range. Even though white matter development is relatively protracted, dense sampling during early childhood old is crucial for approximating these developmental trajectories. The PING sample does include subjects aged 3–21, but the sampling distribution is not uniform over this range, and younger subjects are less represented.

8.5.3 Site effects

Although the PING study has many strengths, one fairly basic limitation is that it is a multisite study of development, and yet participants are not matched on age between sites. As we saw with the ADHD–200 sample, it is also very common to have large site/scanner effects on the imaging data. These two factors combine to make interpretation of all aspects of the data more challenging. For example, considering the age distributions in Figure 8.2, if we saw lower FA in the UCSD subjects, relative to the KKI subjects, is this due to a simple scanner effect (in which case it is a nuisance effect that we would like to control for)? — Or is it due to the UCSD subjects being younger (in which case it is part of the age effect that we are actually trying to study)? One common approach to address this situation would be to simply fit a linear model with both age and site predictors at the same time. However, due to the nonlinearities in the developmental trajectory — the very topic we are trying to study — this method is not perfect. Another way to estimate the site effect would be to analyze the subset of subjects who *do* match on age across sites. Considering nonlinearities, slicing out a cross-section of the data like this is a stronger way to estimate the main site effects. However, this comes with its own challenges, as a matched subset may not always be available. The PING study is in the gray area here, as all sites do span the central age range of late childhood; however, some sites have much sparser sampling (*e.g.* UC Davis, UMMS). The origin of this problem is also important to consider for its implication on designing multisite studies. When the different sites were enrolled, there were no particular quotas on different subgroups of ages. Therefore the sites that began scanning earlier naturally enrolled more older subjects, who are easier to schedule and scan. Fortunately, this problem was noticed part way through enrollment, so the organizers were able to partially correct the problem. Nevertheless it is an important lesson that if we are trying to recruit subjects aged 3–7 from our collaborators, we better be more specific than just giving an age range of 3–21 years. A last thing to always consider in multisite studies is the likely possibility of site:* interactions. Maybe one scanner is less comfortable particularly for younger subjects. Perhaps another scanner has relatively lower sensitivity in older subjects. These effects are not being estimated directly in this study, but they should still be considered when interpreting

multisite results as their effects are likely present in subtle ways.

8.5.4 Future work

The PING study is a huge initiative, and we have only just begun to scratch the surface of the data. Focusing first on the preliminary study of white matter developmental timing we have just described, there are several important areas of future work. First, with such a large sample, and the added diversity that comes with a multisite design, this effort will likely benefit from several more cycles of protocol optimization and quality control. Also, our comparisons of patterns in developmental timing along tracts have thus far been qualitative in nature. Therefore another important step will be to extend these observations with statistical inference; for example, answering the question, "How confident are we that the interesting focal gender effect two thirds of the way along the L IFO in Figure 8.11 is a real phenomenon, and not something that could likely occur by chance?" After these methodological details are flushed out, the rich genomic and behavioral data on these subjects provide many long-term opportunities for scientific discovery. Integrating genomic data will increase power by explaining variance components in our data that are under strong genetic control (Chiang *et al.*, 2011). Additionally, we can generate detailed along-tract maps of the neurobiological correlates of disease candidate genes. Even more directly, we can investigate correlations between behavioral measures and developmental timing. For example, does higher intelligence also correlate with more dynamic development (*i.e.* an earlier time constant) in the white matter, like it does in the cortex, and also show a similar pattern of frontal/temporal localization (Shaw *et al.*, 2006a)?

8.5.5 Conclusions

Although preliminary, this effort is already one of the largest diffusion MRI study to date of white matter development, and one that spans the crucial age range of early childhood. Further, using cutting-edge techniques for along-tract analysis, it is the first time that regional variations in de-

velopmental timing have been mapped within major white matter tracts. These initial observations are contributing to our growing understanding that the patterns of white matter development in the human brain are more complex than previously appreciated, and also inform the broad range of studies that seek to examine disease within the background context of this expected brain maturation. Finally, the wealth of accompanying data on these subjects holds major promise for extending these analyses to other modalities in the future.

Bibliography

1. Adams E., Chau V., Poskitt K. J., Grunau R. E., Synnes A., and Miller S. P. (2010). Tractography-based quantitation of corticospinal tract development in premature newborns. *Journal of Pediatrics* 156(6), 882–8. [DOI]

2. Allen J. S., Damasio H., Grabowski T. J., Bruss J., and Zhang W. (2003). Sexual dimorphism and asymmetries in the gray-white composition of the human cerebrum. *Neuroimage* 18(4), 880–94. [HTML]

3. Ambroise C. and McLachlan G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Sciences of the United States of America* 99(10), 6562–6566.

4. Asato M. R., Terwilliger R., Woo J., and Luna B. (2010). White matter development in adolescence: A DTI study. *Cerebral Cortex* 20(9), 2122–31. [DOI] [HTML]

5. Ashburner J., Csernansky J. G., Davatzikos C., Fox N. C., Frisoni G. B., and Thompson P. M. (2003). Computer-assisted imaging to assess brain structure in healthy and diseased brains. *Lancet Neurology* 2(2), 79–88. [HTML]

6. Ashburner J. and Friston K. J. (2000). Voxel-based morphometry–the methods. *Neuroimage* 11(6 Pt 1), 805–21. [DOI] [HTML]

7. Ashtari M., Cervellione K. L., Hasan K. M., Wu J., McIlree C., Kester H., Ardekani B. A., Roofeh D., Szeszko P. R., and Kumra S. (2007). White matter development during late adolescence in healthy males: a cross-sectional diffusion tensor imaging study. *Neuroimage* 35(2), 501–10. [DOI] [HTML]

8. Astley S. J. (2004). *Diagnostic guide for fetal alcohol spectrum disorders: the 4-digit diagnostic code*. Seattle, WA: University of Washington.

9. Bakken T. E., Roddey J. C., Djurovic S., Akshoomoff N., Amaral D. G., Bloss C. S., Casey B. J., Chang L., Ernst T. M., Gruen J. R., Jernigan T. L., Kaufmann W. E., Kenet T., Kennedy D. N., Kuperman J. M., Murray S. S., Sowell E. R., Rimol L. M., Mattingsdal M., Melle I., Agartz I., Andreassen O. A., Schork N. J., and Dale A. M. (2012). Association of Common Genetic Variants in GPCPD1 with Scaling of Visual Cortical Surface Area in Humans. *Proceedings of the National Academy of Sciences* 109(10), 3985–3990. [DOI] [HTML]

10. Barkovich A. J. (2006). MR imaging of the neonatal brain. *Neuroimaging Clinics of North America* 16(1), 117–35, viii–ix. [DOI] [HTML]

11. Barkovich A. J., Kjos B. O., Jackson D. E., and Norman D. (1988). Normal maturation of the neonatal and infant brain: MR imaging at 1.5 T. *Radiology* 166(1 Pt 1), 173–80. [HTML]

12. Bashat D. B., Kronfeld-Duenias V., Zachor D. A., Ekstein P. M., Hendler T., Tarrasch R., Even A., Levy Y., and Sira L. B. (2007). Accelerated maturation of white matter in young children with autism: a high b value DWI study. *Neuroimage* 37(1), 40–7. [DOI] [HTML]

13. Bashat D. B., Sira L. B., Graif M., Pianka P., Hendler T., Cohen Y., and Assaf Y. (2005). Normal white matter development from infancy to adulthood: comparing diffusion tensor and high b value diffusion weighted MR images. *J Magn Reson Imaging* 21(5), 503–11. [DOI] [HTML]

14. Basser P. J., Mattiello J., and LeBihan D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal* 66(1), 259–67. [DOI] [HTML]

15. Bava S., Thayer R., Jacobus J., Ward M., Jernigan T. L., and Tapert S. F. (2010). Longitudinal characterization of white matter maturation during adolescence. *Brain Research* 1327, 38–46. [DOI] [HTML]

16. Beaulieu C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine* 15(7-8), 435–55. [DOI] [HTML]

17. Beery K. (1997). *The Beery-Buktenica Developmental Test of Visual-Motor Integration: Administration, Scoring, and Teaching Manual*. 4th. Parsippany, NJ: Modern Curriculum Press.

18. Behrens T. E. J., Johansen-Berg H., Woolrich M. W., Smith S. M., Wheeler-Kingshott C. A. M., Boulby P. A., Barker G. J., Sillery E. L., Sheehan K., Ciccarelli O., Thompson A. J., Brady J. M., and Matthews P. M. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience* 6(7), 750–7. [DOI] [HTML]

19. Bellis M. D. D., Keshavan M. S., Beers S. R., Hall J., Frustaci K., Masalehdan A., Noll J., and Boring A. M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex* 11(6), 552–7. [HTML]

20. Benes F. M., Turtle M., Khan Y., and Farol P. (1994). Myelination of a key relay zone in the hippocampal formation occurs in the human brain during childhood, adolescence, and adulthood. *Archives of General Psychiatry* 51(6), 477–84. [HTML]

21. Bengtsson S. L., Nagy Z., Skare S., Forsman L., Forssberg H., and Ullén F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience* 8(9), 1148–50. [DOI] [HTML]

22. Bihan D. L., Breton E., Lallemand D., Grenier P., Cabanis E., and Laval-Jeantet M. (1986). MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology* 161(2), 401–7. [HTML]

23. Blinkhorn S. (2005). Intelligence: a gender bender. *Nature* 438(7064), 31–2. [DOI] [HTML]

24. Bonekamp D., Nagae L. M., Degaonkar M., Matson M., Abdalla W. M. A., Barker P. B., Mori S., and Horská A. (2007). Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. *Neuroimage* 34(2), 733–42. [DOI] [HTML]

25. Bookheimer S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25, 151–88. [DOI] [HTML]

26. Bramen J. E., Hranilovich J. A., Dahl R. E., Forbes E. E., Chen J. E., Toga A. W., Dinov I. D., Worthman C. M., and Sowell E. R. (2011). Puberty influences medial temporal lobe and cortical gray matter maturation differently in boys than girls matched for sexual maturity. *Cerebral Cortex* 21(3), 636–46.

27. Brown J. A., Terashima K. H., Burggren A. C., Ercoli L. M., Miller K. J., Small G. W., and Bookheimer S. Y. (2011). Brain network local interconnectivity loss in aging APOE-4 allele carriers. *Proceedings of the National Academy of Sciences of the United States of America* 108(51). 20760–20765. [DOI] [HTML]

28. Buchem M. A. v., Steens S. C., Vrooman H. A., Zwinderman A. H., McGowan J. C., Rassek M., and Engelbrecht V. (2001). Global estimation of myelination in the developing brain on the basis of magnetization transfer imaging: a preliminary study. *AJNR. American Journal of Neuroradiology* 22(4), 762–6. [HTML]

29. Bullmore E. and Sporns O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* 10(3), 186–98. [DOI]

30. Burzynska A., Preuschhof C., Bäckman L., Nyberg L., Li S., Lindenberger U., and Heekeren H. (2010). Age-related differences in white-matter microstructure: Region-specific patterns of diffusivity. *Neuroimage* 49(3), 2104–12. [DOI]

31. Carreiras M., Seghier M. L., Baquero S., Estévez A., Lozano A., Devlin J. T., and Price C. J. (2009). An anatomical signature for literacy. *Nature* 461(7266), 983–6. [DOI] [HTML]

32. Casey B. J., Cohen J. D., Jezzard P., Turner R., Noll D. C., Trainor R. J., Giedd J., Kaysen D., Hertz-Pannier L., and Rapoport J. L. (1995). Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *Neuroimage* 2(3), 221–9. [DOI] [HTML]

33. Casey B. J., Epstein J. N., Buhle J., Liston C., Davidson M. C., Tonev S. T., Spicer J., Niogi S., Millner A. J., Reiss A., Garrett A., Hinshaw S. P., Greenhill L. L., Shafritz K. M., Vitolo A., Kotler L. A., Jarrett M. A., and Glover G. (2007). Frontostriatal connectivity and its role in cognitive control in parent-child dyads with ADHD. *The American journal of psychiatry* 164(11), 1729–36. [DOI] [HTML]

34. Casey B. J., Tottenham N., Liston C., and Durston S. (2005). Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn Sci* 9(3), 104–10. [DOI] [HTML]

35. Catani M., Howard R. J., Pajevic S., and Jones D. K. (2002). Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *Neuroimage* 17(1), 77–94. [HTML]

36. Caviness V. S., Kennedy D. N., Richelme C., Rademacher J., and Filipek P. A. (1996). The human brain age 7-11 years: a volumetric analysis based on magnetic resonance images. *Cerebral Cortex* 6(5), 726–36. [HTML]

37. Chang L., Cloak C., Jiang C., Farnham S., Tokeshi B., Buchthal S., Hedemark B., Smith L., and Ernst T. (2009). Altered neurometabolites and motor integration in children exposed to methamphetamine in utero. *Neuroimage* 48(2), 391–7. [DOI]

38. Chang L., Smith L. M., LoPresti C., Yonekura M. L., Kuo J., Walot I., and Ernst T. (2004). Smaller subcortical volumes and cognitive deficits in children with prenatal methamphetamine exposure. *Psychiatry Research* 132(2), 95–106. [DOI] [HTML]

39. Chiang M., McMahon K. L., Zubicaray G. I. de, Martin N. G., Hickie I., Toga A. W., Wright M. J., and Thompson P. M. (2011). Genetics of white matter development: a DTI study of 705 twins and their siblings aged 12 to 29. *NeuroImage* 54(3). 2308–2317. [DOI] [HTML]

40. Choi Y. Y., Shamosh N. A., Cho S. H., DeYoung C. G., Lee M. J., Lee J., Kim S. I., Cho Z., Kim K., Gray J. R., and Lee K. H. (2008). Multiple bases of human intelligence revealed by cortical thickness and neural activation. *The Journal of Neuroscience* 28(41), 10323–9. [DOI] [HTML]

41. Clayden J. D., Jentschke S., Muñoz M., Cooper J. M., Chadwick M. J., Banks T., Clark C. A., and Vargha-Khadem F. (2011). Normative Development of White Matter Tracts: Similarities and Differences in Relation to Age, Gender, and Intelligence. *Cerebral Cortex*. [DOI] [HTML]

42. Cloak C., Ernst T., Fujii L., Hedemark B., and Chang L. (2009). Lower diffusion in white matter of children with prenatal methamphetamine exposure. *Neurology* 72(24), 2068–75. [DOI]

43. Colby J. B., Soderberg L., Lebel C., Dinov I. D., Thompson P. M., and Sowell E. R. (2012). Along-tract statistics allow for enhanced tractography analysis. *Neuroimage* 59(4), 3227–3242. [DOI]

44. Colby J. B., Van Horn J. D., and Sowell E. R. (2011). Quantitative in vivo evidence for broad regional gradients in the timing of white matter maturation during adolescence. *Neuroimage* 54(1), 25–31. [DOI]

45. Coles C. D. and Li Z. (2011). Functional neuroimaging in the examination of effects of prenatal alcohol exposure. *Neuropsychology review* 21(2), 119–132.

46. Connor P. D., Sampson P. D., Bookstein F. L., Barr H. M., and Streissguth A. P. (2000). Direct and indirect effects of prenatal alcohol damage on executive function. *Dev Neuropsychol* 18(3), 331–54. [HTML]

47. Conturo T. E., Lori N. F., Cull T. S., Akbudak E., Snyder A. Z., Shimony J. S., McKinstry R. C., Burton H., and Raichle M. E. (1999). Tracking neuronal fiber pathways in the living human brain. *Proceedings of the National Academy of Sciences of the United States of America* 96(18), 10422–7. [HTML]

48. Cooke B. M. (2006). Steroid-dependent plasticity in the medial amygdala. *Neuroscience* 138(3), 997–1005. [DOI] [HTML]

49. Cooper W. O., Habel L. A., Sox C. M., Chan K. A., Arbogast P. G., Cheetham T. C., Murray K. T., Quinn V. P., Stein C. M., Callahan S. T., Fireman B. H., Fish F. A., Kirshner H. S., O'Duffy A., Connell F. A., and Ray W. A. (2011). ADHD drugs and serious cardiovascular events in children and young adults. *The New England Journal of Medicine* 365(20). 1896–1904. [DOI] [HTML]

50. Corouge I., Fletcher P. T., Joshi S., Gouttard S., and Gerig G. (2006). Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. *Medical image analysis* 10(5), 786–98. [DOI] [HTML]

51. Cortes C. and Vapnik V. (1995). Support-Vector Networks. *Machine Learning* 20, 273–297.

52. Courchesne E., Chisum H. J., Townsend J., Cowles A., Covington J., Egaas B., Harwood M., Hinds S., and Press G. A. (2000). Normal brain development and aging: quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology* 216(3), 672–82. [HTML]

53. Dahl R. E. (2004). Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences* 1021, 1–22. [DOI] [HTML]

54. Damasio H., Grabowski T., Frank R., Galaburda A. M., and Damasio A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science* 264(5162), 1102–5. [HTML]

55. Davis S. W., Dennis N. A., Buchler N. G., White L. E., Madden D. J., and Cabeza R. (2009). Assessing the effects of age on long white matter tracts using diffusion tensor tractography. *Neuroimage* 46(2), 530–41. [HTML]

56. Dekaban A. S. (1978). Changes in brain weights during the span of human life: relation of brain weights to body heights and body weights. *Annals of Neurology* 4(4), 345–56. [DOI] [HTML]

57. *DSM-IV* (2000). *Diagnostic and statistical manual of mental disorders (4th ed., text revision)*. Washington, DC: American Psychiatric Association.

58. Dosenbach N. U. F., Nardos B., Cohen A. L., Fair D. A., Power J. D., Church J. A., Nelson S. M., Wig G. S., Vogel A. C., Lessov-Schlaggar C. N., Barnes K. A., Dubis J. W., Feczko E., Coalson R. S., Pruett J. R., Barch D. M., Petersen S. E., and Schlaggar B. L. (2010). Prediction of individual brain maturity using fMRI. *Science* 329(5997), 1358–1361.

59. Draganski B., Gaser C., Busch V., Schuierer G., Bogdahn U., and May A. (2004). Neuroplasticity: changes in grey matter induced by training. *Nature* 427(6972), 311–2. [DOI] [HTML]

60. Duan K., Rajapakse J. C., Wang H., and Azuaje F. (2005). Multiple SVM-RFE for gene selection in cancer classification with expression data. *IEEE transactions on nanobioscience* 4(3). 228–234.

61. Durston S., Pol H. E. H., Casey B. J., Giedd J. N., Buitelaar J. K., and Engeland H. v. (2001). Anatomical MRI of the developing human brain: what have we learned? *Journal of the American Academy of Child and Adolescent Psychiatry* 40(9), 1012–20. [DOI] [HTML]

62. Ecker C., Marquand A., Mourão-Miranda J., Johnston P., Daly E. M., Brammer M. J., Maltezos S., Murphy C. M., Robertson D., Williams S. C., and Murphy D. G. M. (2010). Describing the brain in autism in five dimensions–magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *J Neurosci* 30(32), 10612–23. [DOI] [HTML]

63. Elia J., Borcherding B. G., Rapoport J. L., and Keysor C. S. (1991). Methylphenidate and dextroamphetamine treatments of hyperactivity: are there true nonresponders? *Psychiatry Research* 36(2). 141–155. [HTML]

64. Eliez S. and Reiss A. L. (2000). MRI neuroimaging of childhood psychiatric disorders: a selective review. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 41(6), 679–94. [HTML]

65. Engelbrecht V., Rassek M., Preiss S., Wald C., and Mödder U. (1998). Age-dependent changes in magnetization transfer contrast of white matter in the pediatric brain. *AJNR. American Journal of Neuroradiology* 19(10), 1923–9. [HTML]

66. Engelbrecht V., Scherer A., Rassek M., Witsack H. J., and Mödder U. (2002). Diffusion-weighted MR imaging in the brain in children: findings in the normal brain and in the brain with white matter diseases. *Radiology* 222(2), 410–8.

67. Ethen M. K., Ramadhani T. A., Scheuerle A. E., Canfield M. A., Wyszynski D. F., Druschel C. M., Romitti P. A., and Study N. B. D. P. (2009). Alcohol consumption by women before and during pregnancy. *Maternal and child health journal* 13(2), 274–85. [DOI] [HTML]

68. Fair D. A., Cohen A. L., Dosenbach N. U. F., Church J. A., Miezin F. M., Barch D. M., Raichle M. E., Petersen S. E., and Schlaggar B. L. (2008). The maturing architecture of the brain's default network. *Proc Natl Acad Sci USA* 105(10), 4028–4032.

69. Fischl B. and Dale A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America* 97(20), 11050–5. [DOI] [HTML]

70. Fjell A. M., Westlye L. T., Greve D. N., Fischl B., Benner T., Kouwe A. J. W. v. d., Salat D., Bjørnerud A., Due-Tønnessen P., and Walhovd K. B. (2008). The relationship between diffusion tensor imaging and volumetry as measures of white matter properties. *Neuroimage* 42(4), 1654–68. [DOI] [HTML]

71. Frangou S., Chitins X., and Williams S. C. R. (2004). Mapping IQ and gray matter density in healthy young people. *Neuroimage* 23(3), 800–5. [DOI] [HTML]

72. Froehlich T. E., Lanphear B. P., Epstein J. N., Barbaresi W. J., Katusic S. K., and Kahn R. S. (2007). Prevalence, recognition, and treatment of attention-deficit/hyperactivity disorder in a national sample of US children. *Archives of Pediatrics & Adolescent Medicine* 161(9). 857–864. [DOI] [HTML]

73. Fryer S. L., McGee C. L., Matt G. E., Riley E. P., and Mattson S. N. (2007). Evaluation of psychopathological conditions in children with heavy prenatal alcohol exposure. *Pediatrics* 119(3), e733–41. [DOI] [HTML]

74. Fryer S. L., Schweinsburg B. C., Bjorkquist O. A., Frank L. R., Mattson S. N., Spadoni A. D., and Riley E. P. (2009). Characterization of white matter microstructure in fetal alcohol spectrum disorders. *Alcohol Clin Exp Res* 33(3), 514–21. [DOI] [HTML]

75. Fuster J. M. (2000). Executive frontal functions. *Exp Brain Res* 133(1), 66–70. [HTML]

76. Fuster J. M. (2001). The prefrontal cortex–an update: time is of the essence. *Neuron* 30(2), 319–33. [HTML]

77. Fuster J. M. (2002). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe*. New York: Lippincott-Raven.

78. Galea L. A. M., Spritzer M. D., Barker J. M., and Pawluski J. L. (2006). Gonadal hormone modulation of hippocampal neurogenesis in the adult. *Hippocampus* 16(3), 225–32. [DOI] [HTML]

79. Geschwind N. and Levitsky W. (1968). Human brain: left-right asymmetries in temporal speech region. *Science* 161(837), 186–7. [HTML]

80. Giedd J. N., Blumenthal J., Jeffries N. O., Castellanos F. X., Liu H., Zijdenbos A., Paus T., Evans A. C., and Rapoport J. L. (1999a). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience* 2(10), 861–3. [DOI] [HTML]

81. Giedd J. N., Blumenthal J., Jeffries N. O., Rajapakse J. C., Vaituzis A. C., Liu H., Berry Y. C., Tobin M., Nelson J., and Castellanos F. X. (1999b). Development of the human corpus callosum during childhood and adolescence: a longitudinal MRI study. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 23(4), 571–88.

82. Giedd J. N., Castellanos F. X., Rajapakse J. C., Vaituzis A. C., and Rapoport J. L. (1997). Sexual dimorphism of the developing human brain. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 21(8), 1185–201. [HTML]

83. Giedd J. N., Clasen L. S., Lenroot R., Greenstein D., Wallace G. L., Ordaz S., Molloy E. A., Blumenthal J. D., Tossell J. W., Stayer C., Samango-Sprouse C. A., Shen D., Davatzikos C., Merke D., and Chrousos G. P. (2006). Puberty-related influences on brain development. *Molecular and Cellular Endocrinology* 254-255, 154–62. [DOI] [HTML]

84. Giedd J. N., Rumsey J. M., Castellanos F. X., Rajapakse J. C., Kaysen D., Vaituzis A. C., Vauss Y. C., Hamburger S. D., and Rapoport J. L. (1996a). A quantitative MRI study of the corpus callosum in children and adolescents. *Brain Research. Developmental Brain Research* 91(2), 274–80. [HTML]

85. Giedd J. N., Snell J. W., Lange N., Rajapakse J. C., Casey B. J., Kozuch P. L., Vaituzis A. C., Vauss Y. C., Hamburger S. D., Kaysen D., and Rapoport J. L. (1996b). Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cerebral Cortex* 6(4), 551–60. [HTML]

86. Giedd J. N., Vaituzis A. C., Hamburger S. D., Lange N., Rajapakse J. C., Kaysen D., Vauss Y. C., and Rapoport J. L. (1996c). Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: ages 4-18 years. *The Journal of Comparative Neurology* 366(2), 223–30. [DOI] [HTML]

87. Giorgio A., Watkins K. E., Chadwick M., James S., Winmill L., Douaud G., Stefano N. D., Matthews P. M., Smith S. M., Johansen-Berg H., and James A. C. (2010). Longitudinal changes in grey and white matter during adolescence. *Neuroimage* 49(1), 94–103. [DOI] [HTML]

88. Giorgio A., Watkins K. E., Douaud G., James A. C., James S., Stefano N. D., Matthews P. M., Smith S. M., and Johansen-Berg H. (2008). Changes in white matter microstructure during adolescence. *Neuroimage* 39(1), 52–61. [DOI] [HTML]

89. Gogtay N., Giedd J. N., Lusk L., Hayashi K. M., Greenstein D., Vaituzis A. C., Nugent T. F., Herman D. H., Clasen L. S., Toga A. W., Rapoport J. L., and Thompson P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America* 101(21), 8174–9. [DOI] [HTML]

90. Goldman L. S., Genel M., Bezman R. J., and Slanetz P. J. (1998). Diagnosis and treatment of attention-deficit/hyperactivity disorder in children and adolescents. Council on Scientific Affairs, American Medical Association. *JAMA: The Journal of the American Medical Association* 279(14). 1100–1107. [HTML]

91. Goldstein J. M., Seidman L. J., Horton N. J., Makris N., Kennedy D. N., Caviness V. S., Faraone S. V., and Tsuang M. T. (2001). Normal sexual dimorphism of the adult human brain assessed by in vivo magnetic resonance imaging. *Cereb Cortex* 11(6), 490–7. [HTML]

92. Gould S. J. (1978). Morton's ranking of races by cranial capacity. Unconscious manipulation of data may be a scientific norm. *Science* 200(4341), 503–9.

93. Gould S. J. (1981). *The Mismeasure of Man*. New York: W. W. Norton.

94. Gray R., Mukherjee R. A. S., and Rutter M. (2009). Alcohol consumption during pregnancy and its effects on neurodevelopment: what is known and what remains uncertain. *Addiction* 104(8), 1270–3. [DOI] [HTML]

95. Guerri C., Bazinet A., and Riley E. P. (2009). Foetal Alcohol Spectrum Disorders and alterations in brain and behaviour. *Alcohol and alcoholism* 44(2), 108–14. [DOI] [HTML]

96. Gur R. C., Gunning-Dixon F., Bilker W. B., and Gur R. E. (2002). Sex differences in temporolimbic and frontal brain volumes of healthy adults. *Cerebral Cortex* 12(9), 998–1003. [HTML]

97. Guyon I., Weston J., Barnhill S., and Vapnik V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning* 46(1/3). 389–422.

98. Habel L. A., Cooper W. O., Sox C. M., Chan K. A., Fireman B. H., Arbogast P. G., Cheetham T. C., Quinn V. P., Dublin S., Boudreau D. M., Andrade S. E., Pawloski P. A., Raebel M. A., Smith D. H., Achacoso N., Uratsu C., Go A. S., Sidney S., Nguyen-Huynh M. N., Ray W. A., and Selby J. V. (2011). ADHD medications and risk of serious cardiovascular events in young and middle-aged adults. *JAMA: The Journal of the American Medical Association* 306(24). 2673–2683. [DOI] [HTML]

99. Haier R. J., Jung R. E., Yeo R. A., Head K., and Alkire M. T. (2004). Structural brain variation and general intelligence. *Neuroimage* 23(1), 425–33. [DOI] [HTML]

100. Haier R. J., Jung R. E., Yeo R. A., Head K., and Alkire M. T. (2005). The neuroanatomy of general intelligence: sex matters. *Neuroimage* 25(1), 320–7. [DOI] [HTML]

101. Haug H. (1987). Brain sizes, surfaces, and neuronal sizes of the cortex cerebri: a stereological investigation of man and his variability and a comparison with some mammals (primates, whales, marsupials, insectivores, and one elephant). *The American Journal of Anatomy* 180(2), 126–42. [DOI]

102. Hedges L. V. and Nowell A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science* 269(5220), 41–5. [HTML]

103. Heneghan C. (2010). Why autism can't be diagnosed with brain scans. *the Guardian*. [HTML]

104. Hermoye L., Saint-Martin C., Cosnard G., Lee S., Kim J., Nassogne M., Menten R., Clapuyt P., Donohue P. K., Hua K., Wakana S., Jiang H., Zijl P. C. M. v., and Mori S. (2006). Pediatric diffusion tensor imaging: normal database and observation of the white matter maturation in early childhood. *Neuroimage* 29(2), 493–504. [DOI] [HTML]

105. Herting M. M., Maxwell E. C., Irvine C., and Nagel B. J. (2011). The Impact of Sex, Puberty, and Hormones on White Matter Microstructure in Adolescents. *Cerebral Cortex*. [DOI] [HTML]

106. Howell K. K., Lynch M. E., Platzman K. A., Smith G. H., and Coles C. D. (2006). Prenatal alcohol exposure and ability, academic achievement, and school functioning in adolescence: a longitudinal follow-up. *J Pediatr Psychol* 31(1), 116–26. [DOI] [HTML]

107. Hoyme H. E., May P. A., Kalberg W. O., Kodituwakku P., Gossage J. P., Trujillo P. M., Buckley D. G., Miller J. H., Aragon A. S., Khaole N., Viljoen D. L., Jones K. L., and Robinson L. K. (2005). A practical clinical approach to diagnosis of fetal alcohol spectrum disorders: clarification of the 1996 institute of medicine criteria. *Pediatrics* 115(1), 39–47. [DOI] [HTML]

108. Hsu J., Hecke W. V., Bai C., Lee C., Tsai Y., Chiu H., Jaw F., Hsu C., Leu J., Chen W., and Leemans A. (2010). Microstructural white matter changes in normal aging: A diffusion tensor imaging study with higher-order polynomial regression models. *Neuroimage* 49(1), 32–43. [DOI] [HTML]

109. Huang H. (2010). Structure of the Fetal Brain: What We Are Learning from Diffusion Tensor Imaging. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry* 16(6), 634–49. [DOI]

110. Huang H., Xue R., Zhang J., Ren T., Richards L., Yarowsky P., Miller M., and Mori S. (2009). Anatomical Characterization of Human Fetal Brain Development with Diffusion Tensor Magnetic Resonance Imaging. *The Journal of Neuroscience* 29(13), 4263–4273. [DOI] [HTML]

111. Huang H., Zhang J., Wakana S., Zhang W., Ren T., Richards L. J., Yarowsky P., Donohue P., Graham E., Zijl P. C. M. v., and Mori S. (2006). White and gray matter development in human fetal, newborn and pediatric brains. *Neuroimage* 33(1), 27–38. [DOI]

112. Hüppi P. S., Maier S. E., Peled S., Zientara G. P., Barnes P. D., Jolesz F. A., and Volpe J. J. (1998). Microstructural development of human newborn cerebral white matter assessed in vivo by diffusion tensor magnetic resonance imaging. *Pediatric Research* 44(4), 584–90. [HTML]

113. Huttenlocher P. R. (1979). Synaptic density in human frontal cortex - developmental changes and effects of aging. *Brain Research* 163, 195–205.

114. Huttenlocher P. R. and Dabholkar A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology* 387(2), 167–78. [HTML]

115. Huttenlocher P. R., De Courten C., and Garey L. J. (1982). Synaptic development in human cerebral cortex. *International Journal of Neurology* 17, 144–154.

116. Im K., Lee J., Lee J., Shin Y., Kim I. Y., Kwon J. S., and Kim S. I. (2006). Gender difference analysis of cortical thickness in healthy young adults with surface-based methods. *Neuroimage* 31(1), 31–8. [DOI] [HTML]

117. MMWR (2010). Increasing prevalence of parent-reported attention-deficit/hyperactivity disorder among children — United States, 2003 and 2007. *MMWR. Morbidity and Mortality Weekly Report* 59(44). 1439–1443. [HTML]

118. Irwing P. and Lynn R. (2006). Intelligence: is there a sex difference in IQ scores? *Nature* 442(7098), E1; discussion E1–2. [DOI] [HTML]

119. Jernigan T. L. and Tallal P. (1990). Late childhood changes in brain morphology observable with MRI. *Developmental Medicine and Child Neurology* 32(5), 379–85. [HTML]

120. Jernigan T. L., Trauner D. A., Hesselink J. R., and Tallal P. A. (1991). Maturation of human cerebrum observed in vivo during adolescence. *Brain* 114(5), 2037–49. [HTML]

121. Johannes M., Brase J. C., Fröhlich H., Gade S., Gehrmann M., Fälth M., Sültmann H., and Beissbarth T. (2010). Integration of pathway knowledge into a reweighted recursive feature elimination approach for risk stratification of cancer patients. *Bioinformatics* 26(17). 2136–2144. [DOI] [HTML]

122. Johansen-Berg H. and Rushworth M. F. S. (2009). Using diffusion imaging to study human connectional anatomy. *Annu Rev Neurosci* 32, 75–94. [DOI] [HTML]

123. Jones K. L., Smith D. W., Ulleland C. N., and Streissguth P. (1973). Pattern of malformation in offspring of chronic alcoholic mothers. *Lancet* 1(7815), 1267–71. [HTML]

124. Jorm A. F., Anstey K. J., Christensen H., and Rodgers B. (2004). Gender differences in cognitive abilities: The mediating role of health state and health habits. *Intelligence* 32(1), 7–23. [DOI] [HTML]

125. Kaes T. (1907). *Die Grosshirnrinde des Menschen in ihern Massen und in iherm Fasergehalt*. Jena, Germany: Gustav Fischer.

126. Kandel E. R., Schwartz J. H., and Jessell T. M. (2000). *Principles of Neural Science*. New York: McGraw-Hill.

127. Karama S., Ad-Dab'bagh Y., Haier R., Deary I. J., Lyttelton O., Lepage C., and Evans A. (2009). Positive association between cognitive ability and cortical thickness in a representative US sample of healthy 6 to 18 year-olds. *Intelligence* 37(2), 145–155. [DOI] [HTML]

128. Kemper T. L. (1994). Neuroanatomical and neuropathological changes during aging. In: *Clinical Neurology of Aging*. Ed. by M. L. Albert and J. E. Knoefel. 2nd. New York, NY: Oxford University Press, 8–9.

129. Kim J. H., Budde M. D., Liang H., Klein R. S., Russell J. H., Cross A. H., and Song S. (2006). Detecting axon damage in spinal cord from a mouse model of multiple sclerosis. *Neurobiology of disease* 21(3). 626–32. [DOI] [HTML]

130. Kimura D. (1996). Sex, sexual orientation and sex hormones influence human cognitive function. *Curr Opin Neurobiol* 6(2), 259–63.

131. Klingberg T., Vaidya C. J., Gabrieli J. D., Moseley M. E., and Hedehus M. (1999). Myelination and organization of the frontal white matter in children: a diffusion tensor MRI study. *Neuroreport* 10(13), 2817–21. [HTML]

132. Kortte K. B., Horner M. D., and Windham W. K. (2002). The trail making test, part B: cognitive flexibility or ability to maintain set? *Applied neuropsychology* 9(2), 106–9. [HTML]

133. Kumar A., Sundaram S. K., Sivaswamy L., Behen M. E., Makki M. I., Ager J., Janisse J., Chugani H. T., and Chugani D. C. (2010). Alterations in frontal lobe tracts and corpus callosum in young children with autism spectrum disorder. *Cerebral Cortex* 20(9), 2103–13. [DOI] [HTML]

134. Lagasse L. L., Wouldes T., Newman E., Smith L. M., Shah R. Z., Derauf C., Huestis M. A., Arria A. M., Grotta S. D., Wilcox T., and Lester B. M. (2011). Prenatal methamphetamine exposure and neonatal neurobehavioral outcome in the USA and New Zealand. *Neurotoxicology and teratology* 33(1), 166–75. [DOI] [HTML]

135. Lancaster J. L., Andrews T., Hardies L. J., Dodd S., and Fox P. T. (2003). Three-pool model of white matter. *Journal of Magnetic Resonance Imaging : JMRI* 17(1), 1–10. [DOI] [HTML]

136. Larson K., Russ S. A., Kahn R. S., and Halfon N. (2011). Patterns of comorbidity, functioning, and service use for US children with ADHD, 2007. *Pediatrics* 127(3). 462–470. [DOI] [HTML]

137. Lebel C. and Beaulieu C. (2009). Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Human Brain Mapping* 30(11), 3563–73. [DOI]

138. Lebel C. and Beaulieu C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31(30). 10937–10947. [DOI] [HTML]

139. Lebel C., Caverhill-Godkewitsch S., and Beaulieu C. (2010). Age-related regional variations of the corpus callosum identified by diffusion tensor tractography. *Neuroimage* 52(1), 20–31. [DOI]

140. Lebel C., Gee M., Camicioli R., Wieler M., Martin W., and Beaulieu C. (2012). Diffusion tensor imaging of white matter tract evolution over the lifespan. *NeuroImage* 60(1). 340–352. [DOI] [HTML]

141. Lebel C., Rasmussen C., Wyper K., Walker L., Andrew G., Yager J., and Beaulieu C. (2008a). Brain diffusion abnormalities in children with fetal alcohol spectrum disorder. *Alcoholism, clinical and experimental research* 32(10), 1732–1740. [DOI]

142. Lebel C., Roussotte F., and Sowell E. R. (2011). Imaging the impact of prenatal alcohol exposure on the structure of the developing human brain. *Neuropsychology review* 21(2), 102–118.

143. Lebel C., Walker L., Leemans A., Phillips L., and Beaulieu C. (2008b). Microstructural maturation of the human brain from childhood to adulthood. *Neuroimage* 40(3), 1044–1055. [DOI] [HTML]

144. Le Bihan D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews. Neuroscience* 4(6), 469–80. [DOI] [HTML]

145. Lenroot R. K. and Giedd J. N. (2010). Sex differences in the adolescent brain. *Brain and Cognition* 72(1), 46–55. [DOI] [HTML]

146. Lenroot R. K., Gogtay N., Greenstein D. K., Wells E. M., Wallace G. L., Clasen L. S., Blumenthal J. D., Lerch J., Zijdenbos A. P., Evans A. C., Thompson P. M., and Giedd J. N. (2007). Sexual dimorphism of brain developmental trajectories during childhood and adolescence. *Neuroimage* 36(4), 1065–73. [DOI] [HTML]

147. Li L., Coles C. D., Lynch M. E., and Hu X. (2009). Voxelwise and skeleton-based region of interest analysis of fetal alcohol syndrome and fetal alcohol spectrum disorders in young adults. *Human brain mapping* 30(10), 3265–74. [DOI] [HTML]

148. Liston C., Malter Cohen M., Teslovich T., Levenson D., and Casey B. J. (2011). Atypical prefrontal connectivity in attention-deficit/hyperactivity disorder: pathway to disease or pathological end point? *Biological Psychiatry* 69(12). 1168–1177. [DOI] [HTML]

149. Liston C., Watts R., Tottenham N., Davidson M. C., Niogi S., Ulug A. M., and Casey B. J. (2006). Frontostriatal microstructure modulates efficient recruitment of cognitive control. *Cereb Cortex* 16(4), 553–60. [DOI] [HTML]

150. Liu Y., Balériaux D., Kavec M., Metens T., Absil J., Denolin V., Pardou A., Avni F., Bogaert P. V., and Aeby A. (2010). Structural asymmetries in motor and language networks in a population of healthy preterm neonates at term equivalent age: a diffusion tensor imaging and probabilistic tractography study. *Neuroimage* 51(2), 783–8. [DOI] [HTML]

151. Löbel U., Sedlacik J., Güllmar D., Kaiser W. A., Reichenbach J. R., and Mentzel H. (2009). Diffusion tensor imaging: the normal evolution of ADC, RA, FA, and eigenvalues studied in multiple anatomical regions of the brain. *Neuroradiology* 51(4), 253–63. [DOI] [HTML]

152. Luders E., Narr K. L., Thompson P. M., Rex D. E., Jancke L., Steinmetz H., and Toga A. W. (2004). Gender differences in cortical complexity. *Nature Neuroscience* 7(8), 799–800. [DOI] [HTML]

153. Lu L. H., Johnson A., O'Hare E. D., Bookheimer S. Y., Smith L. M., O'Connor M. J., and Sowell E. R. (2009). Effects of prenatal methamphetamine exposure on verbal memory revealed with functional magnetic resonance imaging. *Journal of developmental and behavioral pediatrics: JDBP* 30(3), 185–92. [DOI] [HTML]

154. Lu L., Leonard C., Thompson P., Kan E., Jolley J., Welcome S., Toga A., and Sowell E. (2007). Normal developmental changes in inferior frontal gray matter are associated with improvement in phonological processing: a longitudinal MRI analysis. *Cerebral Cortex* 17(5), 1092–9. [DOI] [HTML]

155. Luna B., Garver K. E., Urban T. A., Lazar N. A., and Sweeney J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development* 75(5), 1357–72. [DOI] [HTML]

156. Luna B., Padmanabhan A., and O'Hearn K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and cognition* 72(1), 101–13. [DOI] [HTML]

157. Lynn R. and Irwing P. (2004). Sex differences on the progressive matrices: A meta-analysis. *Intelligence* 32(5), 481–498. [DOI] [HTML]

158. MacKay A., Whittall K., Adler J., Li D., Paty D., and Graeb D. (1994). In vivo visualization of myelin water in brain by magnetic resonance. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine* 31(6), 673–7.

159. Marsh R., Gerber A. J., and Peterson B. S. (2008). Neuroimaging studies of normal brain development and their relevance for understanding childhood neuropsychiatric disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 47(11), 1233–51. [DOI] [HTML]

160. Mattson S. N., Goodman A. M., Caine C., Delis D. C., and Riley E. P. (1999). Executive functioning in children with heavy prenatal alcohol exposure. *Alcohol Clin Exp Res* 23(11), 1808–15. [HTML]

161. Mattson S. N. and Riley E. P. (1998). A review of the neurobehavioral deficits in children with fetal alcohol syndrome or prenatal exposure to alcohol. *Alcohol Clin Exp Res* 22(2), 279–94. [HTML]

162. Mattson S. N., Riley E. P., Gramling L., Delis D. C., and Jones K. L. (1998). Neuropsychological comparison of alcohol-exposed children with or without physical features of fetal alcohol syndrome. *Neuropsychology* 12(1), 146–53.

163. Mattson S. N., Roesch S. C., Fagerlund A., Autti-Rämö I., Jones K. L., May P. A., Adnams C. M., Konovalova V., and Riley E. P. (2010). Toward a neurobehavioral profile of fetal alcohol spectrum disorders. *Alcohol Clin Exp Res*. [DOI] [HTML]

164. Ma X., Coles C. D., Lynch M. E., Laconte S. M., Zurkiya O., Wang D., and Hu X. (2005). Evaluation of corpus callosum anisotropy in young adults with fetal alcohol syndrome according to diffusion tensor imaging. *Alcohol Clin Exp Res* 29(7), 1214–22. [HTML]

165. Maxwell J. C. and Rutkowski B. A. (2008). The prevalence of methamphetamine and amphetamine abuse in North America: a review of the indicators, 1992-2007. *Drug and alcohol review* 27(3), 229–35. [DOI] [HTML]

166. May P. A. and Gossage J. P. (2001). Estimating the prevalence of fetal alcohol syndrome. A summary. *Alcohol research and health* 25(3), 159–67. [HTML]

167. May P. A., Gossage J. P., Marais A., Hendricks L. S., Snell C. L., Tabachnick B. G., Stellavato C., Buckley D. G., Brooke L. E., and Viljoen D. L. (2008). Maternal risk factors for fetal alcohol syndrome and partial fetal alcohol syndrome in South Africa: a third study. *Alcohol Clin Exp Res* 32(5), 738–53. [DOI] [HTML]

168. McCann U. D., Wong D. F., Yokoi F., Villemagne V., Dannals R. F., and Ricaurte G. A. (1998). Reduced striatal dopamine transporter density in abstinent methamphetamine and methcathinone users: evidence from positron emission tomography studies with [11C]WIN-35,428. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 18(20), 8417–22. [HTML]

169. McKetin R., Kozel N., Douglas J., Ali R., Vicknasingam B., Lund J., and Li J. (2008). The rise of methamphetamine in Southeast and East Asia. *Drug and alcohol review* 27(3), 220–8. [DOI] [HTML]

170. McKinstry R. C., Mathur A., Miller J. H., Ozcan A., Snyder A. Z., Schefft G. L., Almli C. R., Shiran S. I., Conturo T. E., and Neil J. J. (2002). Radial organization of developing preterm human cerebral cortex revealed by non-invasive water diffusion anisotropy MRI. *Cerebral Cortex* 12(12), 1237–43. [HTML]

171. Merikangas K. R., He J., Brody D., Fisher P. W., Bourdon K., and Koretz D. S. (2010). Prevalence and treatment of mental disorders among US children in the 2001-2004 NHANES. *Pediatrics* 125(1). 75–81. [DOI] [HTML]

172. Michaelis L., Menten M. L., Johnson K. A., and Goody R. S. (2011). The original Michaelis constant: translation of the 1913 Michaelis-Menten paper. *Biochemistry* 50(39). 8264–8269. [DOI] [HTML]

173. Miller S. P. and Ferriero D. M. (2009). From selective vulnerability to connectivity: insights from newborn brain imaging. *Trends in Neurosciences* 32(9), 496–505. [DOI] [HTML]

174. Miyake A., Friedman N. P., Emerson M. J., Witzki A. H., Howerter A., and Wager T. D. (2000). The unity and diversity of executive functions and their contributions to complex ”Frontal Lobe” tasks: a latent variable analysis. *Cogn Psychol* 41(1), 49–100. [DOI] [HTML]

175. Mori S., Crain B. J., Chacko V. P., and Zijl P. C. V. (1999). Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* 45(2), 265–9.

176. Mori S., Oishi K., Jiang H., Jiang L., Li X., Akhter K., Hua K., Faria A. V., Mahmood A., Woods R., Toga A. W., Pike G. B., Neto P. R., Evans A., Zhang J., Huang H., Miller M. I., Zijl P. v., and Mazziotta J. (2008). Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *Neuroimage* 40(2), 570–82. [DOI] [HTML]

177. Morriss M. C., Zimmerman R. A., Bilaniuk L. T., Hunter J. V., and Haselgrove J. C. (1999). Changes in brain water diffusion during childhood. *Neuroradiology* 41(12), 929–34. [HTML]

178. Mukherjee P., Miller J. H., Shimony J. S., Conturo T. E., Lee B. C., Almli C. R., and McKinstry R. C. (2001). Normal brain maturation during childhood: developmental trends characterized with diffusion-tensor MR imaging. *Radiology* 221(2), 349–58. [HTML]

179. Nagy Z., Westerberg H., and Klingberg T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience* 16(7), 1227–33. [DOI] [HTML]

180. Narr K. L., Woods R. P., Thompson P. M., Szeszko P., Robinson D., Dimtcheva T., Gurbani M., Toga A. W., and Bilder R. M. (2007). Relationships between IQ and regional cortical gray matter thickness in healthy adults. *Cerebral Cortex* 17(9), 2163–71. [DOI] [HTML]

181. Neil J. J., Shiran S. I., McKinstry R. C., Schefft G. L., Snyder A. Z., Almli C. R., Akbudak E., Aronovitz J. A., Miller J. P., Lee B. C., and Conturo T. E. (1998). Normal brain in human newborns: apparent diffusion coefficient and diffusion anisotropy measured by using diffusion tensor MR imaging. *Radiology* 209(1), 57–66. [HTML]

182. Neisser U., Boodoo G., Jr T. B., Boykin A., Brody N., Ceci S., Halpern D., Loehlin J., Perloff R., Sternberg R., and Urbina S. (1996). Intelligence: Knowns and Unknowns. *The American Psychologist* 51(2). 77–101. [HTML]

183. Neufang S., Specht K., Hausmann M., Güntürkün O., Herpertz-Dahlmann B., Fink G. R., and Konrad K. (2009). Sex differences and the impact of steroid hormones on the developing human brain. *Cerebral Cortex* 19(2), 464–73. [DOI] [HTML]

184. Nguyen D., Smith L. M., Lagasse L. L., Derauf C., Grant P., Shah R., Arria A., Huestis M. A., Haning W., Strauss A., Grotta S. D., Liu J., and Lester B. M. (2010). Intrauterine growth of infants exposed to prenatal methamphetamine: results from the infant development, environment, and lifestyle study. *J Pediatr* 157(2), 337–9. [DOI] [HTML]

185. Nichols T. E. and Holmes A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping* 15(1), 1–25. [HTML]

186. Nomura Y., Sakuma H., Takeda K., Tagami T., Okuda Y., and Nakagawa T. (1994). Diffusional anisotropy of the human brain assessed with diffusion-weighted MR: relation with normal brain development and aging. *AJNR American Journal of Neuroradiology* 15(2), 231–8. [HTML]

187. Nopoulos P., Flaum M., O'Leary D., and Andreasen N. C. (2000). Sexual dimorphism in the human brain: evaluation of tissue volume, tissue composition and surface anatomy using magnetic resonance imaging. *Psychiatry Research* 98(1), 1–13. [HTML]

188. Norman A. L., Crocker N., Mattson S. N., and Riley E. P. (2009). Neuroimaging and fetal alcohol spectrum disorders. *Developmental disabilities research reviews* 15(3), 209–17. [DOI] [HTML]

189. O'Connor M. J., Frankel F., Paley B., Schonfeld A. M., Carpenter E., Laugeson E. A., and Marquardt R. (2006). A controlled social skills training for children with fetal alcohol spectrum disorders. *Journal of consulting and clinical psychology* 74(4), 639–48. [DOI] [HTML]

190. O'Donnell S., Noseworthy M. D., Levine B., and Dennis M. (2005). Cortical thickness of the frontopolar area in typically developing children and adolescents. *Neuroimage* 24(4), 948–54. [DOI] [HTML]

191. Ostby Y., Tamnes C. K., Fjell A. M., Westlye L. T., Due-Tønnessen P., and Walhovd K. B. (2009). Heterogeneity in subcortical brain development: A structural magnetic resonance imaging study of brain maturation from 8 to 30 years. *The Journal of Neuroscience* 29(38), 11772–82. [DOI] [HTML]

192. Pakkenberg B. and Gundersen H. J. (1997). Neocortical neuron number in humans: effect of sex and age. *The Journal of Comparative Neurology* 384(2), 312–20.

193. Panigrahy A. and Blüml S. (2009). Neuroimaging of pediatric brain tumors: from basic to advanced magnetic resonance imaging (MRI). *J Child Neurol* 24(11), 1343–65. [DOI] [HTML]

194. Paus T., Collins D. L., Evans A. C., Leonard G., Pike B., and Zijdenbos A. (2001). Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain Research Bulletin* 54(3), 255–66. [HTML]

195. Paus T., Nawaz-Khan I., Leonard G., Perron M., Pike G. B., Pitiot A., Richer L., Susman E., Veillette S., and Pausova Z. (2010). Sexual dimorphism in the adolescent brain: Role of testosterone and androgen receptor in global and local volumes of grey and white matter. *Horm Behav* 57(1), 63–75. [DOI] [HTML]

196. Paus T., Zijdenbos A., Worsley K., Collins D. L., Blumenthal J., Giedd J. N., Rapoport J. L., and Evans A. C. (1999). Structural maturation of neural pathways in children and adolescents: in vivo study. *Science* 283(5409), 1908–11. [HTML]

197. Pelham W. E., Foster E. M., and Robb J. A. (2007). The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. *Journal of Pediatric Psychology* 32(6). 711–727. [DOI] [HTML]

198. Peper J. S., Brouwer R. M., Schnack H. G., Baal G. C. M. V., Leeuwen M. v., Berg S. M. v. d., Waal H. A. D. d., Janke A. L., Collins D. L., Evans A. C., Boomsma D. I., Kahn R. S., and Pol H. E. H. (2008). Cerebral white matter in early puberty is associated with luteinizing hormone concentrations. *Psychoneuroendocrinology* 33(7), 909–15. [DOI] [HTML]

199. Peper J. S., Brouwer R. M., Schnack H. G., Baal G. C. v., Leeuwen M. v., Berg S. M. v. d., Waal H. A. D. d., Boomsma D. I., Kahn R. S., and Pol H. E. H. (2009a). Sex steroids and brain structure in pubertal boys and girls. *Psychoneuroendocrinology* 34(3), 332–42. [DOI] [HTML]

200. Peper J. S., Schnack H. G., Brouwer R. M., Baal G. C. M. V., Pjetri E., Székely E., Leeuwen M. v., Berg S. M. v. d., Collins D. L., Evans A. C., Boomsma D. I., Kahn R. S., and Pol H. E. H. (2009b). Heritability of regional and global brain structure at the onset of puberty: a magnetic resonance imaging study in 9-year-old twin pairs. *Human Brain Mapping* 30(7), 2184–96. [DOI] [HTML]

201. Perrin J. S., Hervé P., Leonard G., Perron M., Pike G. B., Pitiot A., Richer L., Veillette S., Pausova Z., and Paus T. (2008). Growth of white matter in the adolescent brain: role of testosterone and androgen receptor. *J Neurosci* 28(38), 9519–24. [DOI] [HTML]

202. Peters M., Jäncke L., Staiger J. F., Schlaug G., Huang Y., and Steinmetz H. (1998). Unsolved problems in comparing brain sizes in Homo sapiens. *Brain and Cognition* 37(2), 254–85. [DOI] [HTML]

203. Pfefferbaum A., Mathalon D. H., Sullivan E. V., Rawles J. M., Zipursky R. B., and Lim K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Archives of Neurology* 51(9), 874–87. [HTML]

204. Phillips O., Nuechterlein K., Clark K., Hamilton L., Asarnow R., Hageman N., Toga A., and Narr K. (2009). Fiber tractography reveals disruption of temporal lobe white matter tracts in schizophrenia. *Schizophrenia Research* 107(1), 30–8. [DOI] [HTML]

205. Pierpaoli C., Jezzard P., Basser P. J., Barnett A., and Chiro G. D. (1996). Diffusion tensor MR imaging of the human brain. *Radiology* 201(3), 637–48. [HTML]

206. Post R. M. and Weiss S. R. B. (1997). Emergent properties of neural systems: How focal molecular neurobiological alterations can affect behavior. *Development and Psychopathology* 9(04), 907–929. [DOI] [HTML]

207. Power J. D., Fair D. A., Schlaggar B. L., and Petersen S. E. (2010). The development of human functional brain networks. *Neuron* 67(5), 735–48. [DOI]

208. Power T. J., Costigan T. E., Leff S. S., Eiraldi R. B., and Landau S. (2001). Assessing ADHD across settings: contributions of behavioral assessment to categorical decision making. *Journal of Clinical Child Psychology* 30(3). 399–412. [DOI] [HTML]

209. Prager A. and Roychowdhury S. (2007). Magnetic resonance imaging of the neonatal brain. *Indian Journal of Pediatrics* 74(2), 173–84. [HTML]

210. Pujol J., López-Sala A., Deus J., Cardoner N., Sebastián-Gallés N., Conesa G., and Capdevila A. (2002). The lateral asymmetry of the human brain studied by volumetric magnetic resonance imaging. *Neuroimage* 17(2), 670–9. [HTML]

211. Pujol J., Vendrell P., Junqué C., Martí-Vilalta J. L., and Capdevila A. (1993). When does human brain development end? Evidence of corpus callosum growth up to adulthood. *Annals of Neurology* 34(1), 71–5. [DOI]

212. Qiu D., Tan H., Zhou K., and Khong P. (2008). Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: Voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. *Neuroimage* 41(2), 223–32. [DOI] [HTML]

213. Rabinowicz T., Petetot J. M., Gartside P. S., Sheyn D., Sheyn T., and de C. (2002). Structure of the cerebral cortex in men and women. *Journal of Neuropathology and Experimental Neurology* 61(1), 46–57. [HTML]

214. Rasmussen C. (2005). Executive functioning and working memory in fetal alcohol spectrum disorder. *Alcohol Clin Exp Res* 29(8), 1359–67. [HTML]

215. Reiss A. L., Abrams M. T., Singer H. S., Ross J. L., and Denckla M. B. (1996). Brain development, gender and IQ in children. A volumetric imaging study. *Brain* 119(5), 1763–74. [HTML]

216. *SAMHSA* (2008). *Results from the 2007 NSDUH: National Findings*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

217. Rex D. E., Ma J. Q., and Toga A. W. (2003). The LONI Pipeline processing environment. *Neuroimage* 19(3), 1033–48. [HTML]

218. Riddle W. R., Donlevy S. C., and Lee H. (2010). Modeling brain tissue volumes over the lifespan: quantitative analysis of postmortem weights and in vivo MR images. *Magnetic Resonance Imaging* 28(5), 716–20. [DOI]

219. Rilling J. K., Glasser M. F., Preuss T. M., Ma X., Zhao T., Hu X., and Behrens T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience* 11(4), 426–8. [DOI] [HTML]

220. Romine C. B. and Reynolds C. R. (2005). A model of the development of frontal lobe functioning: findings from a meta-analysis. *Applied neuropsychology* 12(4), 190–201. [DOI] [HTML]

221. Roussotte F. F., Bramen J. E., Nunez S. C., Quandt L. C., Smith L., O'Connor M. J., Bookheimer S. Y., and Sowell E. R. (2011). Abnormal brain activation during working memory in children with prenatal exposure to drugs of abuse: the effects of methamphetamine, alcohol, and polydrug exposure. *Neuroimage* 54(4), 3067–75. [DOI] [HTML]

222. Roussotte F., Soderberg L., and Sowell E. (2010). Structural, metabolic, and functional brain abnormalities as a result of prenatal exposure to drugs of abuse: evidence from neuroimaging. *Neuropsychol Rev* 20(4), 376–97. [DOI] [HTML]

223. Sakai K. (2008). Task set and prefrontal cortex. *Annu Rev Neurosci* 31, 219–45. [DOI] [HTML]

224. Salat D. H., Tuch D. S., Hevelone N. D., Fischl B., Corkin S., Rosas H. D., and Dale A. M. (2005). Age-related changes in prefrontal white matter measured by diffusion tensor imaging. *Annals of the New York Academy of Sciences* 1064, 37–49. [DOI] [HTML]

225. Sampson P. D., Streissguth A. P., Bookstein F. L., Little R. E., Clarren S. K., Dehaene P., Hanson J. W., and Graham J. M. (1997). Incidence of fetal alcohol syndrome and prevalence of alcohol-related neurodevelopmental disorder. *Teratology* 56(5), 317–26. [DOI] [HTML]

226. Sarkey S., Azcoitia I. n., Garcia-Segura L. M., Garcia-Ovejero D., and DonCarlos L. L. (2008). Classical androgen receptors in non-classical sites in the brain. *Hormones and Behavior* 53(5), 753–64. [DOI] [HTML]

227. Schmithorst V. J., Holland S. K., and Dardzinski B. J. (2008). Developmental differences in white matter architecture between boys and girls. *Human Brain Mapping* 29(6), 696–710. [DOI] [HTML]

228. Schmithorst V. J., Wilke M., Dardzinski B. J., and Holland S. K. (2002). Correlation of white matter diffusivity and anisotropy with age during childhood and adolescence: a cross-sectional diffusion-tensor MR imaging study. *Radiology* 222(1), 212–8. [HTML]

229. Schmithorst V. and Yuan W. (2010). White matter development during adolescence as shown by diffusion MRI. *Brain and cognition* 72(1), 16–25. [DOI] [HTML]

230. Schneider J. F. L., Il'yasov K. A., Hennig J., and Martin E. (2004). Fast quantitative diffusion-tensor imaging of cerebral white matter from the neonatal period to adolescence. *Neuroradiology* 46(4), 258–66. [DOI] [HTML]

231. Schneiderman J. S., Buchsbaum M. S., Haznedar M. M., Hazlett E. A., Brickman A. M., Shihabuddin L., Brand J. G., Torosjan Y., Newmark R. E., Tang C., Aronowitz J., Paul-Odouard R., Byne W., and Hof P. R. (2007). Diffusion tensor anisotropy in adolescents and adults. *Neuropsychobiology* 55(2), 96–111. [DOI] [HTML]

232. Scholz J., Klein M. C., Behrens T. E. J., and Johansen-Berg H. (2009). Training induces changes in white-matter architecture. *Nature Neuroscience* 12(11), 1367–8. [DOI] [HTML]

233. Shaw P. (2007). Intelligence and the developing human brain. *Bioessays* 29(10), 962–73. [DOI] [HTML]

234. Shaw P., Eckstrand K., Sharp W., Blumenthal J., Lerch J. P., Greenstein D., Clasen L., Evans A., Giedd J., and Rapoport J. L. (2007). Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proceedings of the National Academy of Sciences of the United States of America* 104(49). 19649–19654. [DOI] [HTML]

235. Shaw P., Greenstein D., Lerch J., Clasen L., Lenroot R., Gogtay N., EVANS A., Rapoport J., and Giedd J. (2006a). Intellectual ability and cortical development in children and adolescents. *Nature* 440(7084), 676–9. [DOI] [HTML]

236. Shaw P., Kabani N. J., Lerch J. P., Eckstrand K., Lenroot R., Gogtay N., Greenstein D., Clasen L., Evans A., Rapoport J. L., Giedd J. N., and Wise S. P. (2008). Neurodevelopmental trajectories of the human cerebral cortex. *The Journal of Neuroscience* 28(14), 3586–94. [DOI] [HTML]

237. Shaw P., Lerch J., Greenstein D., Sharp W., Clasen L., Evans A., Giedd J., Castellanos F. X., and Rapoport J. (2006b). Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry* 63(5). 540–549. [DOI] [HTML]

238. Shi P., Ray S., Zhu Q., and Kon M. A. (2011). Top scoring pairs for feature selection in machine learning and applications to cancer outcome prediction. *BMC Bioinformatics* 12. 375. [DOI] [HTML]

239. Silk T., Vance A., Rinehart N., Bradshaw J., and Cunnington R. (2008). White-matter abnormalities in attention deficit hyperactivity disorder: A diffusion tensor imaging study. *Human brain mapping* 30(9), 2757–65. [DOI]

240. Skranes J., Lohaugen G. C., Martinussen M., Indredavik M. S., Dale A. M., Haraldseth O., Vangberg T. R., and Brubakk A. (2009). White matter abnormalities and executive function in children with very low birth weight. *Neuroreport* 20(3), 263–6. [HTML]

241. Skranes J., Vangberg T. R., Kulseng S., Indredavik M. S., Evensen K. A. I., Martinussen M., Dale A. M., Haraldseth O., and Brubakk A. (2007). Clinical findings and white matter abnormalities seen on diffusion tensor imaging in adolescents with very low birth weight. *Brain* 130(Pt 3). 654–66. [DOI] [HTML]

242. Slotkin T. A. (1998). Fetal nicotine or cocaine exposure: which one is worse? *The Journal of pharmacology and experimental therapeutics* 285(3), 931–45. [HTML]

243. Smith L. M., Chang L., Yonekura M. L., Grob C., Osborn D., and Ernst T. (2001). Brain proton magnetic resonance spectroscopy in children exposed to methamphetamine in utero. *Neurology* 57(2), 255–60. [HTML]

244. Smith L. M., LaGasse L. L., Derauf C., Grant P., Shah R., Arria A., Huestis M., Haning W., Strauss A., Grotta S. D., Fallone M., Liu J., and Lester B. M. (2008). Prenatal methamphetamine use and neonatal neurobehavioral outcome. *Neurotoxicology and teratology* 30(1), 20–8. [DOI] [HTML]

245. Smith L. M., LaGasse L. L., Derauf C., Grant P., Shah R., Arria A., Huestis M., Haning W., Strauss A., Grotta S. D., Liu J., and Lester B. M. (2006a). The infant development, environment, and lifestyle study: effects of prenatal methamphetamine exposure, polydrug exposure, and poverty on intrauterine growth. *Pediatrics* 118(3), 1149–56. [DOI] [HTML]

246. Smith S. M., Jenkinson M., Johansen-Berg H., Rueckert D., Nichols T. E., Mackay C. E., Watkins K. E., Ciccarelli O., Cader M. Z., Matthews P. M., and Behrens T. E. J. (2006b). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31(4), 1487–505. [DOI] [HTML]

247. Smith S. M., Jenkinson M., Woolrich M. W., Beckmann C. F., Behrens T. E. J., Johansen-Berg H., Bannister P. R., Luca M. D., Drobnjak I., Flitney D. E., Niazy R. K., Saunders J., Vickers J., Zhang Y., Stefano N. D., Brady J. M., and Matthews P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, S208–19. [DOI] [HTML]

248. Smith S. M., Johansen-Berg H., Jenkinson M., Rueckert D., Nichols T. E., Miller K. L., Robson M. D., Jones D. K., Klein J. C., Bartsch A. J., and Behrens T. E. J. (2007). Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. *Nature Protocols* 2(3), 499–503. [DOI] [HTML]

249. Smith S. and Nichols T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1), 83–98. [DOI]

250. Snook L., Paulson L., Roy D., Phillips L., and Beaulieu C. (2005). Diffusion tensor imaging of neurodevelopment in children and young adults. *Neuroimage* 26(4), 1164–73. [DOI] [HTML]

251. Snook L., Plewes C., and Beaulieu C. (2007). Voxel based versus region of interest analysis in diffusion tensor imaging of neurodevelopment. *Neuroimage* 34(1), 243–52. [DOI] [HTML]

252. Song S., Sun S., Ramsbottom M. J., Chang C., Russell J., and Cross A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage* 17(3). 1429–36. [HTML]

253. Sowell E. R., Delis D., Stiles J., and Jernigan T. L. (2001a). Improved memory functioning and frontal lobe maturation between childhood and adolescence: a structural MRI study. *J. Int. Neuropsychol. Soc.* 7(3), 312–22.

254. Sowell E. R., Johnson A., Kan E., Lu L. H., Van Horn J. D., Toga A. W., O'Connor M. J., and Bookheimer S. Y. (2008). Mapping white matter integrity and neurobehavioral correlates in children with fetal alcohol spectrum disorders. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28(6), 1313–9. [DOI] [HTML]

255. Sowell E. R., Leow A. D., Bookheimer S. Y., Smith L. M., O'Connor M. J., Kan E., Rosso C., Houston S., Dinov I. D., and Thompson P. M. (2010). Differentiating prenatal exposure to methamphetamine and alcohol versus alcohol and not methamphetamine using tensor-based brain morphometry and discriminant analysis. *J Neurosci* 30(11), 3876–85. [DOI] [HTML]

256. Sowell E. R., Peterson B. S., Kan E., Woods R. P., Yoshii J., Bansal R., Xu D., Zhu H., Thompson P. M., and Toga A. W. (2007). Sex differences in cortical thickness mapped in 176 healthy individuals between 7 and 87 years of age. *Cerebral Cortex* 17(7), 1550–60. [DOI] [HTML]

257. Sowell E. R., Peterson B. S., Thompson P. M., Welcome S. E., Henkenius A. L., and Toga A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience* 6(3), 309–15. [DOI] [HTML]

258. Sowell E. R., Thompson P. M., Holmes C. J., Batth R., Jernigan T. L., and Toga A. W. (1999a). Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *Neuroimage* 9(6 Pt 1), 587–97. [DOI] [HTML]

259. Sowell E. R., Thompson P. M., Holmes C. J., Jernigan T. L., and Toga A. W. (1999b). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience* 2(10), 859–61. [DOI] [HTML]

260. Sowell E. R., Thompson P. M., Leonard C. M., Welcome S. E., Kan E., and Toga A. W. (2004a). Longitudinal mapping of cortical thickness and brain growth in normal children. *The Journal of Neuroscience* 24(38), 8223–31. [DOI] [HTML]

261. Sowell E. R., Thompson P. M., Rex D., Kornsand D., Tessner K. D., Jernigan T. L., and Toga A. W. (2002a). Mapping sulcal pattern asymmetry and local cortical surface gray matter distribution in vivo: maturation in perisylvian cortices. *Cerebral Cortex* 12(1), 17–26.

262. Sowell E. R., Thompson P. M., and Tessner K. D. (2001b). Mapping continued brain growth and gray matter density reduction in dorsal frontal cortex: Inverse relationships during postadolescent brain maturation. *The Journal of Neuroscience* 21(22).

263. Sowell E. R., Thompson P. M., and Toga A. W. (2004b). Mapping changes in the human cortex throughout the span of life. *Neuroscientist* 10(4), 372–92.

264. Sowell E. R., Trauner D. A., Gamst A., and Jernigan T. L. (2002b). Development of cortical and subcortical brain structures in childhood and adolescence: a structural MRI study. *Developmental Medicine and Child Neurology* 44(1), 4–16. [HTML]

265. Spadoni A. D., McGee C. L., Fryer S. L., and Riley E. P. (2007). Neuroimaging and fetal alcohol spectrum disorders. *Neuroscience and biobehavioral reviews* 31(2), 239–45. [DOI] [HTML]

266. Spear L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and biobehavioral reviews* 24(4), 417–63. [HTML]

267. Spohr H., Willms J., and Steinhausen H. (2007). Fetal alcohol spectrum disorders in young adulthood. *The Journal of pediatrics* 150(2), 175–9, 179.e1. [DOI] [HTML]

268. Steinberg L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences* 9(2), 69–74. [DOI] [HTML]

269. Streissguth A. P., Aase J. M., Clarren S. K., Randels S. P., LaDue R. A., and Smith D. F. (1991). Fetal alcohol syndrome in adolescents and adults. *JAMA* 265(15), 1961–7. [HTML]

270. Strogatz S. H. (2001). Exploring complex networks. *Nature* 410(6825). 268–276. [DOI] [HTML]

271. Suzuki Y., Matsuzawa H., Kwee I. L., and Nakada T. (2003). Absolute eigenvalue diffusion tensor analysis for human brain maturation. *NMR in Biomedicine* 16(5), 257–60. [DOI] [HTML]

272. Tamnes C., Ostby Y., Fjell A., Westlye L., Due-Tønnessen P., and Walhovd K. (2010). Brain Maturation in Adolescence and Young Adulthood: Regional Age-Related Changes in Cortical Thickness and White Matter Volume and Microstructure. *Cerebral Cortex* 20(3), 534–48. [DOI]

273. Thomas C., Avidan G., Humphreys K., Jung K.-j., Gao F., and Behrmann M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature Neuroscience* 12(1), 29–31. [DOI] [HTML]

274. Thompson B. L., Levitt P., and Stanwood G. D. (2009). Prenatal exposure to drugs: effects on brain development and implications for policy and education. *Nature reviews. Neuroscience* 10(4), 303–312. [DOI] [HTML]

275. Thompson P. M., Cannon T. D., Narr K. L., Erp T. v., Poutanen V. P., Huttunen M., Lönnqvist J., Standertskjöld-Nordenstam C. G., Kaprio J., Khaledy M., Dail R., Zoumalan C. I., and Toga A. W. (2001). Genetic influences on brain structure. *Nature Neuroscience* 4(12), 1253–8. [DOI] [HTML]

276. Thompson P. M., Giedd J. N., Woods R. P., MacDonald D., Evans A. C., and Toga A. W. (2000). Growth patterns in the developing brain detected by using continuum mechanical tensor maps. *Nature* 404(6774), 190–3. [DOI] [HTML]

277. Thompson P. M., Hayashi K. M., Simon S. L., Geaga J. A., Hong M. S., Sui Y., Lee J. Y., Toga A. W., Ling W., and London E. D. (2004). Structural abnormalities in the brains of human subjects who use methamphetamine. *The Journal of neuroscience* 24(26), 6028–6036. [DOI] [HTML]

278. Toga A. W. and Thompson P. M. (2005). Genetics of brain structure and intelligence. *Annual Review of Neuroscience* 28, 1–23. [DOI] [HTML]

279. Tsujimoto S. (2008). The prefrontal cortex: functional neural development during early childhood. *Neuroscientist* 14(4), 345–58. [DOI] [HTML]

280. Tzarouchi L. C., Astrakas L. G., Xydis V., Zikou A., Kosta P., Drougia A., Andronikou S., and Argyropoulou M. I. (2009). Age-related grey matter changes in preterm infants: an MRI study. *Neuroimage* 47(4), 1148–53. [DOI] [HTML]

281. Vafaie H. and Imam I. F. (1994). Feature selection methods: Genetic algorithms vs greedy-like search. *Proceedings Of The International Conference On Fuzzy And Intelligent Control Systems*. [HTML]

282. Vapnik V. and Lerner A. (1963). Pattern recognition using generalized portrait method. *Automation and Remote Control* 24, 774–780.

283. Verhoeven J. S., Sage C. A., Leemans A., Hecke W. V., Callaert D., Peeters R., Cock P. D., Lagae L., and Sunaert S. (2010). Construction of a stereotaxic DTI atlas with full diffusion tensor information for studying white matter maturation from childhood to adolescence using tractography-based segmentations. *Human Brain Mapping* 31(3), 470–86. [DOI] [HTML]

284. Vignaud J. (1966). Development of the nervous system in early life. Part III. Radiological Study of the Normal Skull in Premature and Newborn Infants. In: *Human Development*. Ed. by F. T. Falkner. Philadelphia: WB Saunders Co.

285. Von Economo C. V. (1929). *The Cytoarchitectonics of the Human Cerebral Cortex*. London: Oxford Medical Publications.

286. Wakana S., Caprihan A., Panzenboeck M. M., Fallon J. H., Perry M., Gollub R. L., Hua K., Zhang J., Jiang H., Dubey P., Blitz A., Zijl P. v., and Mori S. (2007). Reproducibility of quantitative tractography methods applied to cerebral white matter. *Neuroimage* 36(3), 630–44. [DOI] [HTML]

287. Wakana S., Jiang H., Nagae-Poetscher L. M., Zijl P. C. M. v., and Mori S. (2004). Fiber tract-based atlas of human white matter anatomy. *Radiology* 230(1), 77–87. [DOI] [HTML]

288. Wechsler D. (2003). *Wechsler Intelligence Scale for Children*. 4th. San Antonio, TX: The Psychological Corporation.

289. Westlye L. T., Walhovd K. B., Dale A. M., Bjørnerud A., Due-Tønnessen P., Engvig A., Grydeland H., Tamnes C. K., Østby Y., and Fjell A. M. (2010). Differentiating maturational and aging-related changes of the cerebral cortex by use of thickness and signal intensity. *Neuroimage* 52(1), 172–85. [DOI]

290. Wilke M., Krägeloh-Mann I., and Holland S. K. (2007). Global and local development of gray and white matter volume in normal children and adolescents. *Experimental Brain Research* 178(3), 296–307. [DOI] [HTML]

291. Wilke M., Sohn J., Byars A. W., and Holland S. K. (2003). Bright spots: correlations of gray matter volume with IQ in a normal pediatric population. *Neuroimage* 20(1), 202–15. [HTML]

292. Willerman L., Schultz R., Rutledge J. N., and Bigler E. D. (1991). In vivo brain size and intelligence. *Intelligence* 15(2), 223–228. [DOI] [HTML]

293. Witelson S. F., Beresh H., and Kigar D. L. (2006). Intelligence and brain size in 100 post-mortem brains: sex, lateralization and age factors. *Brain* 129(Pt 2), 386–98. [DOI] [HTML]

294. Witte A. V., Savli M., Holik A., Kasper S., and Lanzenberger R. (2010). Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain. *Neuroimage* 49(2), 1205–12. [DOI] [HTML]

295. Wolff S. D. and Balaban R. S. (1989). Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. *Magnetic Resonance in Medicine* 10(1), 135–44. [HTML]

296. Wolraich M., Brown L., Brown R. T., DuPaul G., Earls M., Feldman H. M., Ganiats T. G., Kaplanek B., Meyer B., Perrin J., Pierce K., Reiff M., Stein M. T., and Visser S. (2011). ADHD: clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics* 128(5). 1007–1022. [DOI] [HTML]

297. Woolrich M. W., Jbabdi S., Patenaude B., Chappell M., Makni S., Behrens T., Beckmann C., Jenkinson M., and Smith S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage* 45(1 Suppl), S173–86. [DOI] [HTML]

298. Wozniak J. R. and Lim K. O. (2006). Advances in white matter imaging: a review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. *Neuroscience and Biobehavioral Reviews* 30(6), 762–74. [DOI] [HTML]

299. Wozniak J. R., Mueller B. A., Chang N., Muetzel R. L., Caros L., and Lim K. O. (2006). Diffusion tensor imaging in children with fetal alcohol spectrum disorders. *Alcohol Clin Exp Res* 30(10), 1799–806. [DOI] [HTML]

300. Wozniak J. R. and Muetzel R. L. (2011). What does diffusion tensor imaging reveal about the brain and cognition in fetal alcohol spectrum disorders? *Neuropsychology review* 21(2), 133–147.

301. Yakovlev P. I. and Lecours A. R. (1967). The myelogenetic cycles of regional maturation of the brain. In: *Regional Development of the Brain Early in Life*. Ed. by A. Minkowski. Oxford, England: Blackwell Scientific Publications Inc., 3–70.

302. Zang Y., Jiang T., Lu Y., He Y., and Tian L. (2004). Regional homogeneity approach to fMRI data analysis. *NeuroImage* 22(1). 394–400. [DOI] [HTML]

303. Zhang J., Evans A., Hermoye L., Lee S., Wakana S., Zhang W., Donohue P., Miller M. I., Huang H., Wang X., Zijl P. C. M. v., and Mori S. (2007). Evidence of slow maturation of the superior longitudinal fasciculus in early childhood by diffusion tensor imaging. *Neuroimage* 38(2), 239–47. [DOI] [HTML]

304. Zhang J., Konkle A. T. M., Zup S. L., and McCarthy M. M. (2008). Impact of sex and hormones on new cells in the developing rat hippocampus: a novel source of sex dimorphism? *The European Journal of Neuroscience* 27(4), 791–800. [DOI] [HTML]

305. Zhou X. and Tuck D. P. (2007). MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data. *Bioinformatics (Oxford, England)* 23(9). 1106–1114. [DOI] [HTML]

306. Zhu H., Styner M., Tang N., Liu Z., Lin W., and Gilmore J. (2010). FRATS: Functional regression analysis of DTI tract statistics. *IEEE transactions on medical imaging* 29(4), 1039–49. [DOI] [HTML]