Title: White matter microstructure shows sex differences in late childhood: Evidence from 6,797 children

Author List: Katherine E. Lawrence;[1] Zvart Abaryan;[1] Emily Laltoo;[1] Leanna M. Hernandez;[2] Michael Gandal;[2,3,4] James T. McCracken;[2] Paul M. Thompson[1]

Affiliations: [1]Imaging Genetics Center, Mark and Mary Stevens Neuroimaging & Informatics Institute, University of Southern California, Marina del Rey, CA, USA. [2]Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA, USA. [3]Department of Neurology, Center for Autism Research and Treatment, Semel Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [4]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA.

Corresponding Author Information: Katherine E. Lawrence; phone: (323) 442-7246; fax: (323) 442-0137; email: Katherine.Lawrence@ini.usc.edu

Funding Statement: This work was supported by the National Institute of Mental Health (grant numbers R01MH116147 to P.M.T., F32MH122057 to K.E.L., and K00MH119663 to L.M.H.).

Conflict of Interest Disclosure: P.M.T. reports a research grant from Biogen, Inc. for research unrelated to this manuscript. J.T.M. reports consultant income from Roche, TRIS Pharmaceuticals, Octapharma, and GW Pharmaceuticals, expert witness income from Lannett, and research contracts with Roche, Octapharma, and GW Pharmaceuticals for research unrelated to this manuscript. The other authors report no conflicts of interest.

Acknowledgments: This research was conducted using data obtained from the ABCD NIMH Data Archive (NDA) release 3.0 (DOI: 10.15154/1519007).

Abstract

Sex differences in white matter microstructure have been robustly demonstrated in the adult brain using both conventional and advanced diffusion-weighted magnetic resonance imaging (dMRI) approaches. However, the effect of sex on white matter microstructure prior to adulthood remains poorly understood, with previous developmental work focusing on conventional microstructure metrics and yielding mixed results. Here we thoroughly and rigorously characterized sex differences in white matter microstructure among over 6,000 children from the Adolescent Brain Cognitive Development (ABCD) Study who were between 9 and 10 years old. Microstructure was quantified using both the conventional model - diffusion tensor imaging (DTI) - and an advanced model, restriction spectrum imaging (RSI). DTI metrics included fractional anisotropy (FA) and mean, axial, and radial diffusivity (MD, AD, RD). RSI metrics included normalized isotropic, directional, and total intracellular diffusion (N0, ND, NT). We found significant and replicable sex differences in DTI or RSI microstructure metrics in every white matter region examined across the brain. The impact of sex on FA was regionally specific. Across white matter regions, boys exhibited greater MD, AD, and RD than girls, on average. Girls displayed increased N0, ND, and NT compared to boys, on average, suggesting greater cell and neurite density in girls. Together, these robust and replicable findings provide an important foundation for understanding sex differences in health and disease.

Keywords: sex differences, white matter, diffusion-weighted MRI, diffusion tensor imaging, restriction spectrum imaging, microstructure, development

Introduction

Sex differences in white matter microstructure are consistently shown in the adult human brain (Jahanshad & Thompson, 2017; Salminen et al., 2021). Such *in vivo* neuroimaging studies have used both conventional and advanced diffusion-weighted magnetic resonance imaging (dMRI) models to extensively characterize sex differences in large-scale adult samples (Cox et al., 2016; Lawrence et al., 2021; Ritchie et al., 2018). These analyses have firmly established white matter sex differences in the human brain, but we lack a complete understanding of how such differences may manifest prior to adulthood. Previous developmental work examining the impact of participant sex on white matter microstructure has used the conventional dMRI model, diffusion tensor imaging (DTI), and yielded mixed findings (Basser, Mattiello, & LeBihan, 1994a, 1994b; Kaczkurkin, Raznahan, & Satterthwaite, 2019; Tamnes, Roalf, Goddings, & Lebel, 2018). Most such studies have included samples of around 100 youth or less from a range of developmental stages, such as from late childhood to late adolescence (Bava et al., 2011; Clayden et al., 2012; Herting et al., 2017; Herting, Maxwell, Irvine, & Nagel, 2012; Schmithorst, Holland, & Dardzinski, 2008; Seunarine et al., 2016). In one of the larger studies to date, Krogsrud et al. found no significant sex effects in fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), or radial diffusivity (RD) at either timepoint of their longitudinal sample of 159 children aged 4 to 11 years old (Krogsrud et al., 2016). In a longitudinal study of 203 subjects first scanned at 9 years old, 9-year old girls exhibited greater FA in pure white matter compared to their male counterparts (Brouwer et al., 2012). Lopez-Vicente et al. investigated sex differences in 3,031 youth aged 8 to 12 years and found that boys generally exhibited greater MD, AD, and RD than girls, whereas sex differences in FA were more regionally specific (Lopez-Vicente et al., 2021). As a whole, this prior work suggests sex differences in white matter microstructure during development, but the exact nature of these differences remains poorly understood.

Previous developmental studies assessing the impact of participant sex on white matter microstructure have used the conventional dMRI model, diffusion tensor imaging (DTI), and the

resulting microstructure metrics FA, MD, AD, and RD. However, the DTI model is unable to capture complex white matter architecture in the brain, such as crossing or dispersing fibers (Alexander, Lee, Lazar, & Field, 2007; Basser et al., 1994a, 1994b; Jones, 2008). The DTI metrics FA, MD, AD, and RD are also further limited by their lack of specificity in characterizing the microstructural environment. These well-established limitations of DTI are addressed by more advanced dMRI models, including the advanced model, restriction spectrum imaging (RSI). RSI models multiple underlying fiber populations per voxel, thereby allowing for the resolution of complex white matter configurations, and furthermore provides microstructure measures that exhibit greater specificity than DTI (White, Leergaard, D'Arceuil, Bjaalie, & Dale, 2013; White et al., 2014; White, McDonald, et al., 2013).

Here we thoroughly characterize sex differences in white matter microstructure in over 6,000 children from the Adolescent Brain Cognitive Development (ABCD) Study using both the conventional dMRI model, DTI, and the advanced model, RSI (Barch et al., 2018; Casey et al., 2018; Garavan et al., 2018; Hagler et al., 2019; Volkow et al., 2018). We hypothesized that the effect of sex on FA would be regionally dependent and boys would exhibit greater MD, AD, and RD than girls; the association between RSI metrics and sex was an open question due to the lack of comparable previous work. As the impact of participant sex may evolve over development, our analyses specifically focused on a narrow age range in late childhood, such that all participants were between 9 and 10 years old (Clayden et al., 2012; Herting et al., 2017; Schmithorst et al., 2008; Seunarine et al., 2016; Simmonds, Hallquist, Asato, & Luna, 2014; but see Krogsrud et al., 2016; Lebel & Beaulieu, 2011; Palmer et al., 2021). We directly confirmed the reproducibility and robustness of all findings by using split half replication and considering a wide range of potential confounds in our analyses. To the best of our knowledge, this comprehensive study represents the largest investigation to date of white matter microstructure sex differences in a developmental sample.

Material and Methods

Participants

The ABCD Study is a large-scale, on-going longitudinal investigation of brain development with data collected across 21 data acquisition sites (Barch et al., 2018; Casey et al., 2018; Garavan et al., 2018; Volkow et al., 2018). Participating children were recruited from schools within geographical areas that approximate the demographic diversity of the United States. Subjects completed the baseline timepoint between the ages of 9 to 10 years old. Full sampling and exclusion procedures for this epidemiologically-informed sample are described in detail elsewhere (Garavan et al., 2018). The ABCD Study protocol was approved by the centralized Institutional Review Board at the University of California, San Diego, and informed assent and consent were obtained from each participant and their legal guardian.

The current work was conducted using de-identified tabulated neuroimaging, demographic, and behavioral data from the baseline timepoint of the ABCD NIMH Data Archive (NDA) release 3.0 (DOI: 10.15154/1519007). A complete participant flowchart is shown in **Figure S1**. Briefly, subjects were only included in the current analyses if they had complete nuisance covariate data (see *Statistical Analyses*) and complete dMRI data that also passed quality control (Hagler et al., 2019). Intersex participants, or participants whose biological sex assigned at birth did not align with their parent-report gender identity, were excluded from analyses (Mueller et al., 2021). Children were also excluded if they were the sibling of another subject in the study, as statistical models run in all participants using relatedness as a random effect failed to converge; the sibling retained for analyses was selected randomly. Our final sample consisted of 6,797 children between the ages of 9 to 10 years old (48.1% female). To demonstrate the reproducibility of all findings, this final sample was split into a discovery cohort and a replication cohort. The discovery cohort included 3,399 children (47.4% female), and the replication cohort included 3,398 children (48.8% female).

Descriptive statistics are presented for the discovery and replication cohorts in **Table 1.** The reported statistical comparisons were completed in R 3.6.1 using t-tests or chi-squared tests, as appropriate (R Core Team, 2016). The male and female groups were matched on all core demographic variables, including age, household income, parental education, race, and ethnicity (all discovery and replication *p*s > 0.1). As expected, significant sex differences were observed in pubertal development, externalizing behaviors, and intracranial volume. Boys were less pubertally advanced than girls, as assessed by the parent-report Pubertal Development Scale (discovery and replication *p*s < 0.001) (PDS; Petersen, Crockett, Richards, & Boxer, 1988). The male group also exhibited higher measures of externalizing behaviors than girls, as quantified by raw externalizing problem scores from the parent-report Child Behavior Checklist (discovery and replication *p*s < 0.001) (CBCL; Achenbach, 2009). Intracranial volume was significantly larger among boys than girls (discovery and replication *p*s < 0.001). No significant differences were observed between males and females in raw internalizing problem scores from the parent-report CBCL (discovery and replication *p*s > 0.5) or in mean head motion (discovery and replication *p*s > 0.6).

MRI Acquisition and Processing

T1-weighted structural MRI and dMRI scans were acquired as described previously (Casey et al., 2018). Briefly, scans were acquired using a harmonized acquisition protocol across 21 ABCD sites on Siemens Prisma, Philips, or GE 750 3-tesla scanners. The 3D T1-weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) structural MRI scan had isotropic voxel dimensions of 1 mm and was collected without multi-band acceleration. Diffusion-weighted MRI scans consisted of 96 diffusion encoding directions across four diffusion-weighted shells: 6 directions at b = 500 s/mm^2, 15 directions at b = 1000 s/mm^2, 15 directions at b = 2000 s/mm^2, and 60 directions at b = 3000 s/mm^2. Voxel dimensions were 1.7 mm isotropic and a multi-band acceleration factor of 3 was used.

Processing steps for the T1-weighted structural MRI and dMRI scans are detailed extensively elsewhere (Hagler et al., 2019). In brief, T1-weighted images underwent correction for gradient non-linearity distortions using scanner-specific, nonlinear transformations provided by the MRI scanner manufacturers (Jovicich et al., 2006; Wald, Schmitt, & Dale, 2001) and were subsequently registered to standard space (Friston et al., 1995). Preprocessing for the diffusion-weighted MRI scans included eddy current correction, censoring and interpolation of motion-contaminated slices, gradient distortion correction using pairs of reverse-phase encoded b = 0 s/mm^2 dMRI images, and registration to the T1-weighted scan.

Two reconstruction models were fit to the dMRI data: the conventional model, DTI, and the advanced model, RSI (Basser et al., 1994a, 1994b; Hagler et al., 2019; Jones, 2008; White, Leergaard, et al., 2013; White et al., 2014; White, McDonald, et al., 2013). DTI models a single fiber orientation per voxel by fitting a single diffusion tensor, or ellipsoid, for each voxel in the brain (Basser et al., 1994a, 1994b). The DTI model was fit using diffusion-weighted shells equal to or less than 1000 s/mm^2 (inner shell DTI) and, separately, using all collected gradient strengths (full shell DTI) (Hagler et al., 2019). To improve comparability with previous work, our analyses here focus on the inner shell DTI measures (Kaczkurkin et al., 2019; Tamnes et al., 2018). Microstructure metrics derived from DTI included FA, MD, AD, and RD. FA reflects the degree of diffusion anisotropy, and MD represents the average diffusivity in all directions. AD reflects the diffusion parallel to the primary diffusion axis, and RD captures the diffusion perpendicular to the primary diffusion axis. Conventionally, but with several caveats, higher FA is considered to reflect greater white matter integrity or a reduction in crossing fibers (Alexander et al., 2007; Thomason & Thompson, 2011). Higher MD is conventionally thought to indicate increased extracellular volumes or decreased cellular density, lower AD is conventionally believed to reflect axonal pruning or axonal damage, and increased RD is conventionally related to less myelination or the presence of myelin injury (Alexander et al., 2007; Song et al., 2003; Song et al., 2002). However, the DTI model has well-established limitations regarding its inability to resolve complex fiber

configurations and the non-specificity of its microstructure metrics (Alexander et al., 2007; Basser et al., 1994a, 1994b; Jones, 2008).

RSI is an advanced diffusion model that allows for the resolution of complex white matter architecture and provides microstructure indices with greater specificity than DTI (White, Leergaard, et al., 2013; White et al., 2014; White, McDonald, et al., 2013). RSI separately models restricted (intracellular) and hindered (extracellular) diffusion in each voxel using fourth order spherical harmonic functions, allowing for the modeling of multiple diffusion orientations within each voxel. Our analyses here focus on RSI metrics derived from intracellular diffusion, as these may provide more intuitive interpretations than measures calculated from extracellular diffusion by virtue of more directly reflecting diffusion within cell bodies and neurites (White, Leergaard, et al., 2013; White et al., 2014; White, McDonald, et al., 2013). RSI exhibits conceptual similarities to the dMRI model neurite orientation dispersion and density imaging (NODDI), in that they are both advanced multi-compartment models that capture intracellular diffusion (White, Leergaard, et al., 2013; White et al., 2014; White, McDonald, et al., 2013; Zhang, Schneider, Wheeler-Kingshott, & Alexander, 2012); however, NODDI only provides a measure of total intracellular volume fraction, whereas RSI separately captures isotropic and anisotropic intracellular diffusion. Here we analyzed the following intracellular microstructure indices from RSI (Hagler et al., 2019): normalized isotropic (N0), normalized directional (ND) and normalized total (NT) intracellular diffusion. Each RSI metric is defined as the Euclidean norm of the corresponding model coefficients divided by the norm of all model coefficients and is a unitless measure that ranges from 0 to 1. N0 is derived from the 0th order spherical harmonic coefficient and reflects isotropic intracellular diffusion. ND is calculated from the norm of the 2nd and 4th order spherical harmonic coefficients and captures oriented intracellular diffusion. NT is derived from the norm of the 0th, 2nd, and 4th order spherical harmonic coefficients and reflects the overall intracellular diffusion. A number of biological processes may contribute to the intracellular signal fraction, including variability in myelination, astrocytes, and microglia. In addition to the potential impact of these

processes, higher N0 may reflect increased cell density, higher ND may relate to increased neurite density, and greater NT values may indicate an overall larger intracellular space.

Major white matter tracts across the brain were labeled using the probabilistic fiber tract atlas, AtlasTrack (Hagler et al., 2009; Hagler et al., 2019). Mean values for each DTI and RSI metric were extracted for each white matter fiber tract region of interest (ROI), and the following bilateral ROIs were examined in the current study (**Figure S2, Table S1**): all white matter fibers, anterior thalamic radiation (ATR), corpus callosum (CC), cingulate cingulum (CGC), parahippocampal cingulum (CGH), corticospinal/pyramidal tract (CST), *forceps major* (Fmaj), *forceps minor* (Fmin), fornix (FX), fornix excluding the fimbria (FXcut), inferior fronto-occipital fasciculus (IFO), inferior frontal superior frontal cortex (IFSFC), inferior longitudinal fasciculus (ILF), superior corticostriate (SCS), frontal superior corticostriate (fSCS), parietal superior corticostriate (pSCS), striatal inferior frontal cortex (SIFC), superior longitudinal fasciculus (SLF), parietal superior longitudinal fasciculus (pSLF), temporal superior longitudinal fasiculus (tSLF), and uncinate fasciculus (UNC).

Statistical Analyses

Linear mixed models were used to examine microstructural sex differences in white matter fiber tract ROIs across the brain. All regressions were completed in R 3.6.1 using the *lme4* package and included the following nuisance covariates as fixed effects: age, household income, parental education, race, and ethnicity; MRI scanner was modeled as a random effect in all analyses. Results were considered significant and replicable if they survived a 5% false discovery rate (FDR) applied across the number of ROIs in the discovery cohort ($q<0.05$) and demonstrated $p<0.05$ in the replication cohort (Benjamini & Hochberg, 1995; Hernandez et al., 2020). Effect sizes are reported for all analyses as the magnitude of the standardized regression coefficients (standardized betas; βs), reflecting that group comparisons were completed as regressions to allow for the inclusion of nuisance covariates.

Primary analyses characterized the effect of sex on inner shell DTI metrics and intracellular RSI metrics in the bilateral fiber tract ROIs (**Table S1**). Supplementary analyses investigated full shell DTI and extracellular RSI indices for completeness (**Figure S3**). Unless otherwise specified, all subsequent references to DTI correspond to inner shell DTI, and all subsequent references to RSI correspond to intracellular RSI.

We completed a number of analyses to establish the robustness of our primary findings. First, we examined sex differences separately in the left and right hemispheres. Second, we included pubertal development as a covariate to statistically control for known sex differences in pubertal maturation among this age group (Herting et al., 2020). Third, we repeated primary analyses when covarying raw dimensional externalizing or internalizing problem scores from the parent-report CBCL. Fourth, we statistically controlled for mean head motion (Yendiki, Koldewyn, Kakunoori, Kanwisher, & Fischl, 2014). Lastly, intracranial volume was included as a nuisance covariate (Eliot, Ahmed, Khan, & Patel, 2021).

Follow-up analyses expanded on our primary DTI and RSI findings by directly assessing the relative sensitivity of these two models to sex effects. Specifically, we statistically compared the effect size of the most sensitive DTI metric with that of the most sensitive RSI metric for each fiber tract ROI. In line with all other analyses, results were only deemed significant and replicable if they survived FDR correction in the discovery cohort ($q<0.05$) and demonstrated $p<0.05$ in the replication cohort (Hernandez et al., 2020).

Results

Significant and replicable sex differences were observed in DTI or RSI microstructure metrics for every bilateral white matter tract ROI examined (**Figure 1A, Table S2-S8**). The directionality of sex effects on FA varied across white matter ROIs. FA was higher in boys than girls in limbic white matter tract ROIs (CGC, CGH, FX, FXcut) and thalamic projections (ATR) (discovery βs=0.08-0.21; replication βs=0.09-0.20), on average. In contrast, greater FA was

observed among girls in association tract ROIs (IFO, ILF, tSLF, IFSFC, UNC), corticostriatal projections (SCS, fSCS, pSCS, SIFC), and across all white matter fibers on average, compared to boys (discovery βs=0.08-0.25; replication βs=0.07-0.27). Sex effects were highly consistent across ROIs for all other examined microstructure indices: MD, AD, RD, N0, ND, and NT. Boys typically exhibited greater MD, AD, and RD than girls, on average, indicating higher diffusivity in boys (discovery βs=0.12-0.49; replication βs=0.07-0.43). Girls generally displayed higher N0, ND, and NT relative to boys, suggesting greater cell and neurite density in girls (discovery βs=0.08-0.45; replication βs=0.07-0.42). Across microstructure measures, the most substantial sex differences were observed in association tract ROIs (IFO, ILF, tSLF, IFSFC) and superior corticostriatal projections (SCS, fSCS, pSCS). The impact of sex was the least pronounced, albeit still significant and replicable, in limbic white matter tract ROIs (FX, FXcut), commissural tract ROIs (Fmin, Fmaj), and sensorimotor and thalamic projections (CST, ATR).

Follow-up analyses considering sex differences separately in the left and right hemisphere yielded highly similar results to our primary analyses examining bilateral white matter tract ROIs (**Figure 1B-C**, discovery βs=0.05-0.45; replication βs=0.05-0.41). Results were also largely consistent when controlling for pubertal development (**Figure 2**, discovery βs=0.08-0.50; replication βs=0.06-0.40); a small number of ROI and dMRI metric combinations no longer exhibited significant and replicable sex effects, and sex differences were now also observed for FA in commissural tract ROIs (CC, Fmin). Follow-up analyses including dimensional externalizing or internalizing problems as a nuisance covariate provided very similar results to our primary analyses (**Figure 3A-B**, discovery βs=0.05-0.49; replication βs = 0.06-0.43). In sum, our sex difference findings were highly consistent across hemispheres, as well as when statistically controlling for pubertal development, externalizing problems, and internalizing problems.

To further assess the robustness of our sex difference results, we repeated analyses when controlling for head motion and intracranial volume. Results were highly similar when covarying

for head motion (**Figure S4**, discovery βs=0.07-0.49; replication βs=0.06-0.43). Patterns of sex differences were also largely comparable when controlling for intracranial volume (**Figure S5**, discovery βs=0.07-0.42; replication βs=0.05-0.43). Some ROI and dMRI metric combinations no longer exhibited significant and replicable sex effects, or vice versa; a small number of ROI and DTI metrics, but not RSI metrics, exhibited reversed directionality of their findings (**Figure S5**). As a whole, our patterns of observed sex differences were largely consistent when controlling for head motion and intracranial volume.

Lastly, we expanded on our primary analyses by considering the relative sensitivity of DTI and RSI metrics to sex effects. Among DTI metrics, MD or AD generally detected sex differences the most sensitively. N0 typically exhibited the greatest sensitivity among the RSI metrics. When directly contrasting the relative sensitivity of DTI and RSI measures to sex effects, minimal differences were found. The only significant and replicable difference was observed in the ATR, such that the DTI metric AD was more sensitive to sex effects than the RSI metric N0 (discovery FDR=0.015; replication *p*=8.5E-05).

Discussion

Here we rigorously characterized sex differences in white matter microstructure in the largest developmental sample to date using both the conventional dMRI model, DTI, and the advanced model, RSI. Previous work in adults has robustly established the impact of participant sex on adult white matter, as assessed by both conventional and advanced microstructure measures (Jahanshad & Thompson, 2017; Salminen et al., 2021). However, potential sex effects on white matter microstructure prior to adulthood have remained poorly understood, with previous developmental work only including conventional microstructure metrics and yielding mixed results (Kaczkurkin et al., 2019; Tamnes et al., 2018).

In our sample of over 6,000 children between the ages of 9 to 10 years old, we found significant and replicable sex differences in DTI and RSI microstructure measures across the brain. The impact of sex on FA was regionally specific, whereas sex effects on MD, AD, and RD were generally consistent across white matter regions, such that boys exhibited greater diffusivity than girls, on average. Girls typically displayed increased N0, ND, and NT compared to boys, on average, suggesting increased cell and neurite density in girls. Given that our results remained consistent when statistically controlling for puberty, the observed effect of sex in late childhood seems unlikely to be driven entirely by the change in sex hormones associated with the onset of puberty (Crone & Dahl, 2012). Such late childhood sex differences may instead be due at least in part to the impact of X chromosome genes and/or the organizational effects of intrauterine testosterone (Lentini, Kasahara, Arver, & Savic, 2013; Mallard et al., 2021; Salminen et al., 2021).

The impact of participant sex was most pronounced in association and superior corticostriatal white matter tract ROIs. Association fibers connect distributed cortical regions within the same hemisphere, and corticostriatal tracts connect cortical areas across the brain with the striatum (Buyanova & Arsalidou, 2021; Hagler et al., 2009). The impact of participant sex was the least notable in commissural and limbic (fornix) white matter tract ROIs, as well as sensorimotor (corticospinal/pyramidal tract) and thalamic (anterior thalamic radiation) projection tract ROIs. Commissural tracts connect the left and right hemispheres, and the fornix connects areas involved in memory (Buyanova & Arsalidou, 2021; Catani & Thiebaut de Schotten, 2008; Hagler et al., 2009). The corticospinal/pyramidal tract connects the motor cortex to the spinal cord, and the anterior thalamic radiation connects the thalamus with the frontal cortex (Buyanova & Arsalidou, 2021; Catani & Thiebaut de Schotten, 2008; Hagler et al., 2009). As association tracts generally exhibit more prolonged development than commissural and projection tracts, our results suggest that sex differences in late childhood may be most substantial in fibers with protracted developmental trajectories (Tamnes et al., 2018).

The effect of sex on white matter was assessed in our sample by applying both the conventional microstructure model, DTI, and the advanced microstructure model, RSI. When directly contrasting the sensitivity of these two models to sex effects, the two approaches were largely comparable. However, the DTI model has well-established limitations regarding its inability to resolve complex white matter architecture and the non-specificity of its microstructure metrics (Alexander et al., 2007; Basser et al., 1994a, 1994b; Jones, 2008). RSI, in comparison, can resolve complex fiber configurations and provides more refined microstructure measures, offering greater insight into the underlying neurobiology than DTI (White, Leergaard, et al., 2013; White et al., 2014; White, McDonald, et al., 2013). Prior work has together indicated that the relative sensitivity of conventional and advanced dMRI metrics may also depend on the specific neurobiology underlying the scientific question of interest, allowing for the possibility that DTI and RSI may exhibit significantly different sensitivity to effects beyond participant sex (Lawrence et al., 2021; Nir et al., 2019; Pines et al., 2020).

Our findings overall expand on prior developmental studies that have examined the impact of sex on FA, MD, AD, or RD in samples of around 200 youth or less and reported conflicting results (Bava et al., 2011; Brouwer et al., 2012; Clayden et al., 2012; Herting et al., 2017; Herting et al., 2012; Krogsrud et al., 2016; Schmithorst et al., 2008; Seunarine et al., 2016). Supporting the importance of large sample sizes in stabilizing results (Button et al., 2013; Marek et al., 2020), the directionality of our sex difference findings are generally consistent with the largest previously published study of developmental sex effects in FA, MD, AD, and RD, which assessed 3,031 youth aged 8 to 12 years old (Lopez-Vicente et al., 2021). Intriguingly, our results contrast with those of previous adult studies, including our own, which robustly characterized sex differences in samples as large as 3,513 to 15,628 adults (Cox et al., 2016; Lawrence et al., 2021; Ritchie et al., 2018). We found that late childhood sex differences in FA were regionally specific, but sex differences in MD, AD, and RD were generally consistent across white matter regions. The opposite pattern is observed in adults: sex differences in FA are generally consistent across tracts

in adulthood, but sex differences in MD, AD, and RD are regionally specific (Cox et al., 2016; Lawrence et al., 2021; Ritchie et al., 2018). Previous work in humans has suggested an association between sex hormones, such as estrogen or testosterone, and white matter microstructure (Herting et al., 2012; Nabulsi et al., 2020; van Hemmen et al., 2017). The impact of estrogen, progesterone, and testosterone on white matter has also been confirmed in animal studies, which have demonstrated that such hormones directly impact processes such as oligodendrocyte proliferation, maturation, and cell death (Cerghet, Skoff, Swamydas, & Bessert, 2009; Schumacher et al., 2012). The observed difference in sex effects between childhood and adulthood may thus be attributable to changing levels of sex hormones across the lifespan, including rising levels of estrogen and testosterone in adolescence due to puberty (Crone & Dahl, 2012); this hypothesis should be directly tested in future samples that range from childhood through adulthood and include hormonal data.

The current study has a number of important strengths, including the unprecedented sample size, the use of both conventional and advanced microstructure models, the narrow age range of the sample, and the demonstrated reproducibility and robustness of our results. Our findings provide an important foundation for future work dissecting the mechanisms driving sex differences in white matter microstructure, including the potential impact of genetic factors, pubertal development, sex hormones, and environmental contributors (Herting et al., 2017; Herting et al., 2012; Mallard et al., 2021; Nabulsi et al., 2020; Salminen et al., 2021). A range of adolescent-onset neuropsychiatric disorders furthermore exhibit sex differences in their prevalence or presentation and display altered white matter microstructure (Favre et al., 2019; Kelly et al., 2018; Meyer & Lee, 2019; Paus, Keshavan, & Giedd, 2008; Salminen et al., 2021; van Velzen et al., 2020). By characterizing white matter sex differences prior to adolescence, the current work provides a critical foundation for understanding the neurobiological mechanisms that may underlie the emergence of these disorders, including potential sex differences in such mechanisms (Meyer & Lee, 2019; Paus et al., 2008; Salminen et al., 2021). Future work should

expand on the current analyses by examining sex differences in younger samples and charting such differences longitudinally, as well as investigating how such differences may relate to behavior (Tamnes et al., 2018; van Eijk et al., 2021). Additionally, further animal and postmortem studies are needed to more precisely delineate the exact neurobiology underpinning sex differences during development.

In conclusion, we thoroughly characterized white matter sex differences in late childhood using both conventional and advanced microstructure measures. Our results demonstrate replicable and robust sex differences in white matter microstructure across the brain in the largest developmental sample to date. These findings provide an important foundation for understanding sex differences in health and disease.

References

Achenbach, T. M. (2009). *The Achenbach System of Empirically Based Assessment (ASEBA): Development, Findings, Theory and Applications*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families.

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics, 4*(3), 316-329. doi:10.1016/j.nurt.2007.05.011

Barch, D. M., Albaugh, M. D., Avenevoli, S., Chang, L., Clark, D. B., Glantz, M. D., . . . Sher, K. J. (2018). Demographic, physical and mental health assessments in the adolescent brain and cognitive development study: Rationale and description. *Dev Cogn Neurosci, 32*, 55-66. doi:10.1016/j.dcn.2017.10.010

Basser, P. J., Mattiello, J., & LeBihan, D. (1994a). Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B, 103*(3), 247-254. doi:10.1006/jmrb.1994.1037

Basser, P. J., Mattiello, J., & Lebihan, D. (1994b). MR Diffusion Tensor Spectroscopy and Imaging. *Biophysical Journal, 66*(1), 259-267. doi:Doi 10.1016/S0006-3495(94)80775-1

Bava, S., Boucquey, V., Goldenberg, D., Thayer, R. E., Ward, M., Jacobus, J., & Tapert, S. F. (2011). Sex differences in adolescent white matter architecture. *Brain Res, 1375*, 41-48. doi:10.1016/j.brainres.2010.12.051

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology, 57*(1), 289-300. Retrieved from <Go to ISI>://WOS:A1995QE45300017

Brouwer, R. M., Mandl, R. C., Schnack, H. G., van Soelen, I. L., van Baal, G. C., Peper, J. S., . . . Hulshoff Pol, H. E. (2012). White matter development in early puberty: a longitudinal volumetric and diffusion tensor imaging twin study. *PLoS One, 7*(4), e32316. doi:10.1371/journal.pone.0032316

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafo, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci, 14*(5), 365-376. doi:10.1038/nrn3475

Buyanova, I. S., & Arsalidou, M. (2021). Cerebral White Matter Myelination and Relations to Age, Gender, and Cognition: A Selective Review. *Front Hum Neurosci, 15*, 662031. doi:10.3389/fnhum.2021.662031

Casey, B. J., Cannonier, T., Conley, M. I., Cohen, A. O., Barch, D. M., Heitzeg, M. M., . . . Workgroup, A. I. A. (2018). The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. *Dev Cogn Neurosci, 32*, 43-54. doi:10.1016/j.dcn.2018.03.001

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex, 44*(8), 1105-1132. doi:10.1016/j.cortex.2008.05.004

Cerghet, M., Skoff, R. P., Swamydas, M., & Bessert, D. (2009). Sexual dimorphism in the white matter of rodents. *J Neurol Sci, 286*(1-2), 76-80. doi:10.1016/j.jns.2009.06.039

Clayden, J. D., Jentschke, S., Munoz, M., Cooper, J. M., Chadwick, M. J., Banks, T., . . . Vargha-Khadem, F. (2012). Normative development of white matter tracts: similarities and differences in relation to age, gender, and intelligence. *Cereb Cortex, 22*(8), 1738-1747. doi:10.1093/cercor/bhr243

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., . . . Deary, I. J. (2016). Ageing and brain white matter structure in 3,513 UK Biobank participants. *Nat Commun, 7*, 13629. doi:10.1038/ncomms13629

Crone, E. A., & Dahl, R. E. (2012). Understanding adolescence as a period of social-affective engagement and goal flexibility. *Nat Rev Neurosci, 13*(9), 636-650. doi:10.1038/nrn3313

Eliot, L., Ahmed, A., Khan, H., & Patel, J. (2021). Dump the "dimorphism": Comprehensive synthesis of human brain studies reveals few male-female differences beyond size. *Neurosci Biobehav Rev, 125*, 667-697. doi:10.1016/j.neubiorev.2021.02.026

Favre, P., Pauling, M., Stout, J., Hozer, F., Sarrazin, S., Abe, C., . . . Group, E. B. D. W. (2019). Widespread white matter microstructural abnormalities in bipolar disorder: evidence from mega- and meta-analyses across 3033 individuals. *Neuropsychopharmacology, 44*(13), 2285-2293. doi:10.1038/s41386-019-0485-6

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., & Frackowiak, R. S. J. (1995). Spatial registration and normalization of images. *Human Brain Mapping, 3*(3), 165-189. doi:DOI 10.1002/hbm.460030303

Garavan, H., Bartsch, H., Conway, K., Decastro, A., Goldstein, R. Z., Heeringa, S., . . . Zahs, D. (2018). Recruiting the ABCD sample: Design considerations and procedures. *Dev Cogn Neurosci, 32*, 16-22. doi:10.1016/j.dcn.2018.04.004

Hagler, D. J., Jr., Ahmadi, M. E., Kuperman, J., Holland, D., McDonald, C. R., Halgren, E., & Dale, A. M. (2009). Automated white-matter tractography using a probabilistic diffusion tensor atlas: Application to temporal lobe epilepsy. *Hum Brain Mapp, 30*(5), 1535-1547. doi:10.1002/hbm.20619

Hagler, D. J., Jr., Hatton, S., Cornejo, M. D., Makowski, C., Fair, D. A., Dick, A. S., . . . Dale, A. M. (2019). Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage, 202*, 116091. doi:10.1016/j.neuroimage.2019.116091

Hernandez, L. M., Kim, M., Hernandez, C., Thompson, W., Fan, C. C., Galván, A., . . . Gandal, M. (2020). The genetic etiology of childhood insomnia: Longitudinal gene-brain-behavior associations in the ABCD study. *medRxiv*, 2020.2010.2002.20204735. doi:10.1101/2020.10.02.20204735

Herting, M. M., Kim, R., Uban, K. A., Kan, E., Binley, A., & Sowell, E. R. (2017). Longitudinal changes in pubertal maturation and white matter microstructure. *Psychoneuroendocrinology, 81*, 70-79. doi:10.1016/j.psyneuen.2017.03.017

Herting, M. M., Maxwell, E. C., Irvine, C., & Nagel, B. J. (2012). The impact of sex, puberty, and hormones on white matter microstructure in adolescents. *Cereb Cortex, 22*(9), 1979-1992. doi:10.1093/cercor/bhr246

Herting, M. M., Uban, K. A., Gonzalez, M. R., Baker, F. C., Kan, E. C., Thompson, W. K., . . . Sowell, E. R. (2020). Correspondence Between Perceived Pubertal Development and Hormone Levels in 9-10 Year-Olds From the Adolescent Brain Cognitive Development Study. *Front Endocrinol (Lausanne), 11*, 549928. doi:10.3389/fendo.2020.549928

Jahanshad, N., & Thompson, P. M. (2017). Multimodal neuroimaging of male and female brain structure in health and disease across the life span. *J Neurosci Res, 95*(1-2), 371-379. doi:10.1002/jnr.23919

Jones, D. K. (2008). Studying connections in the living human brain with diffusion MRI. *Cortex, 44*(8), 936-952. doi:10.1016/j.cortex.2008.05.002

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., . . . Dale, A. (2006). Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage, 30*(2), 436-443. doi:10.1016/j.neuroimage.2005.09.046

Kaczkurkin, A. N., Raznahan, A., & Satterthwaite, T. D. (2019). Sex differences in the developing brain: insights from multimodal neuroimaging. *Neuropsychopharmacology, 44*(1), 71-85. doi:10.1038/s41386-018-0111-z

Kelly, S., Jahanshad, N., Zalesky, A., Kochunov, P., Agartz, I., Alloza, C., . . . Donohoe, G. (2018). Widespread white matter microstructural differences in schizophrenia across 4322 individuals: results from the ENIGMA Schizophrenia DTI Working Group. *Mol Psychiatry, 23*(5), 1261-1269. doi:10.1038/mp.2017.170

Krogsrud, S. K., Fjell, A. M., Tamnes, C. K., Grydeland, H., Mork, L., Due-Tonnessen, P., . . . Walhovd, K. B. (2016). Changes in white matter microstructure in the developing brain--

A longitudinal diffusion tensor imaging study of children from 4 to 11 years of age. *Neuroimage, 124*(Pt A), 473-486. doi:10.1016/j.neuroimage.2015.09.017

Lawrence, K. E., Nabulsi, L., Santhalingam, V., Abaryan, Z., Villalon-Reina, J. E., Nir, T. M., . . . Thompson, P. M. (2021). Age and Sex Effects on Advanced White Matter Microstructure Measures in 15,628 Older Adults: A UK Biobank Study. *bioRxiv*, 2020.2009.2018.304345. doi:10.1101/2020.09.18.304345

Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues from childhood into adulthood. *J Neurosci, 31*(30), 10937-10947. doi:10.1523/JNEUROSCI.5302-10.2011

Lentini, E., Kasahara, M., Arver, S., & Savic, I. (2013). Sex differences in the human brain and the impact of sex chromosomes and sex hormones. *Cereb Cortex, 23*(10), 2322-2336. doi:10.1093/cercor/bhs222

Lopez-Vicente, M., Lamballais, S., Louwen, S., Hillegers, M., Tiemeier, H., Muetzel, R. L., & White, T. (2021). White matter microstructure correlates of age, sex, handedness and motor ability in a population-based sample of 3031 school-age children. *Neuroimage, 227*, 117643. doi:10.1016/j.neuroimage.2020.117643

Mallard, T. T., Liu, S., Seidlitz, J., Ma, Z., Moraczewski, D., Thomas, A., & Raznahan, A. (2021). X-chromosome influences on neuroanatomical variation in humans. *Nat Neurosci*. doi:10.1038/s41593-021-00890-w

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., . . . Dosenbach, N. U. F. (2020). Towards Reproducible Brain-Wide Association Studies. *bioRxiv*, 2020.2008.2021.257758. doi:10.1101/2020.08.21.257758

Meyer, H. C., & Lee, F. S. (2019). Translating Developmental Neuroscience to Understand Risk for Psychiatric Disorders. *Am J Psychiatry, 176*(3), 179-185. doi:10.1176/appi.ajp.2019.19010091

Mueller, S. C., Guillamon, A., Zubiaurre-Elorza, L., Junque, C., Gomez-Gil, E., Uribe, C., . . . Luders, E. (2021). The Neuroanatomy of Transgender Identity: Mega-Analytic Findings From the ENIGMA Transgender Persons Working Group. *J Sex Med, 18*(6), 1122-1129. doi:10.1016/j.jsxm.2021.03.079

Nabulsi, L., Lawrence, K. E., Santhalingam, V., Abaryan, Z., Boyle, C. P., Villalon-Reina, J. E., . . . Thompson, P. M. (2020). Exogenous sex hormone effects on brain microstructure in women: a diffusion MRI study in the UK Biobank. *bioRxiv*, 2020.2009.2018.304154. doi:10.1101/2020.09.18.304154

Nir, T. M., Thomopoulos, S. I., Villalon-Reina, J. E., Zavaliangos-Petropulu, A., Dennis, E. L., Reid, R. I., . . . Weiner, M. W. (2019). *Multi-Shell Diffusion MRI Measures of Brain Aging: A Preliminary Comparison From ADNI3.* Paper presented at the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019).

Palmer, C. E., Pecheva, D., Iversen, J., Hagler, D., Sugrue, L., Nedelec, P., . . . Dale, A. M. (2021). Microstructural development across white matter from 9-13 years. *bioRxiv*, 2021.2006.2004.447102. doi:10.1101/2021.06.04.447102

Paus, T., Keshavan, M., & Giedd, J. N. (2008). Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci, 9*(12), 947-957. doi:10.1038/nrn2513

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *J Youth Adolesc, 17*(2), 117-133. doi:10.1007/BF01537962

Pines, A. R., Cieslak, M., Larsen, B., Baum, G. L., Cook, P. A., Adebimpe, A., . . . Satterthwaite, T. D. (2020). Leveraging multi-shell diffusion for studies of brain development in youth and young adulthood. *Dev Cogn Neurosci, 43*, 100788. doi:10.1016/j.dcn.2020.100788

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ritchie, S. J., Cox, S. R., Shen, X., Lombardo, M. V., Reus, L. M., Alloza, C., . . . Deary, I. J. (2018). Sex Differences in the Adult Human Brain: Evidence from 5216 UK Biobank Participants. *Cereb Cortex, 28*(8), 2959-2975. doi:10.1093/cercor/bhy109

Salminen, L. E., Tubi, M. A., Bright, J., Thomopoulos, S. I., Wieand, A., & Thompson, P. M. (2021). Sex is a defining feature of neuroimaging phenotypes in major brain disorders. *Hum Brain Mapp*. doi:10.1002/hbm.25438

Schmithorst, V. J., Holland, S. K., & Dardzinski, B. J. (2008). Developmental differences in white matter architecture between boys and girls. *Hum Brain Mapp, 29*(6), 696-710. doi:10.1002/hbm.20431

Schumacher, M., Hussain, R., Gago, N., Oudinet, J. P., Mattern, C., & Ghoumari, A. M. (2012). Progesterone synthesis in the nervous system: implications for myelination and myelin repair. *Front Neurosci, 6*, 10. doi:10.3389/fnins.2012.00010

Seunarine, K. K., Clayden, J. D., Jentschke, S., Munoz, M., Cooper, J. M., Chadwick, M. J., . . . Clark, C. A. (2016). Sexual Dimorphism in White Matter Developmental Trajectories Using Tract-Based Spatial Statistics. *Brain Connect, 6*(1), 37-47. doi:10.1089/brain.2015.0340

Simmonds, D. J., Hallquist, M. N., Asato, M., & Luna, B. (2014). Developmental stages and sex differences of white matter and behavioral development through adolescence: a longitudinal diffusion tensor imaging (DTI) study. *Neuroimage, 92*, 356-368. doi:10.1016/j.neuroimage.2013.12.044

Song, S. K., Sun, S. W., Ju, W. K., Lin, S. J., Cross, A. H., & Neufeld, A. H. (2003). Diffusion tensor imaging detects and differentiates axon and myelin degeneration in mouse optic nerve after retinal ischemia. *Neuroimage, 20*(3), 1714-1722. doi:10.1016/j.neuroimage.2003.07.005

Song, S. K., Sun, S. W., Ramsbottom, M. J., Chang, C., Russell, J., & Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage, 17*(3), 1429-1436. doi:10.1006/nimg.2002.1267

Tamnes, C. K., Roalf, D. R., Goddings, A. L., & Lebel, C. (2018). Diffusion MRI of white matter microstructure development in childhood and adolescence: Methods, challenges and progress. *Dev Cogn Neurosci, 33*, 161-175. doi:10.1016/j.dcn.2017.12.002

Thomason, M. E., & Thompson, P. M. (2011). Diffusion imaging, white matter, and psychopathology. *Annu Rev Clin Psychol, 7*, 63-85. doi:10.1146/annurev-clinpsy-032210-104507

van Eijk, L., Zhu, D., Couvy-Duchesne, B., Strike, L. T., Lee, A. J., Hansell, N. K., . . . Zietsch, B. P. (2021). Are Sex Differences in Human Brain Structure Associated With Sex Differences in Behavior? *Psychol Sci*, 956797621996664. doi:10.1177/0956797621996664

van Hemmen, J., Saris, I. M. J., Cohen-Kettenis, P. T., Veltman, D. J., Pouwels, P. J. W., & Bakker, J. (2017). Sex Differences in White Matter Microstructure in the Human Brain Predominantly Reflect Differences in Sex Hormone Exposure. *Cereb Cortex, 27*(5), 2994-3001. doi:10.1093/cercor/bhw156

van Velzen, L. S., Kelly, S., Isaev, D., Aleman, A., Aftanas, L. I., Bauer, J., . . . Schmaal, L. (2020). White matter disturbances in major depressive disorder: a coordinated analysis across 20 international cohorts in the ENIGMA MDD working group. *Mol Psychiatry, 25*(7), 1511-1525. doi:10.1038/s41380-019-0477-2

Volkow, N. D., Koob, G. F., Croyle, R. T., Bianchi, D. W., Gordon, J. A., Koroshetz, W. J., . . . Weiss, S. R. B. (2018). The conception of the ABCD study: From substance use to a broad NIH collaboration. *Dev Cogn Neurosci, 32*, 4-7. doi:10.1016/j.dcn.2017.10.002

Wald, L., Schmitt, F., & Dale, A. (2001). Systematic spatial distortion in MRI due to gradient non-linearities. *Neuroimage, 13*(6), S50-S50. Retrieved from <Go to ISI>://WOS:000169106300051

White, N. S., Leergaard, T. B., D'Arceuil, H., Bjaalie, J. G., & Dale, A. M. (2013). Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation. *Hum Brain Mapp, 34*(2), 327-346. doi:10.1002/hbm.21454

White, N. S., McDonald, C., Farid, N., Kuperman, J., Karow, D., Schenker-Ahmed, N. M., . . . Dale, A. M. (2014). Diffusion-weighted imaging in cancer: physical foundations and applications of restriction spectrum imaging. *Cancer Res, 74*(17), 4638-4652. doi:10.1158/0008-5472.CAN-13-3534

White, N. S., McDonald, C. R., Farid, N., Kuperman, J. M., Kesari, S., & Dale, A. M. (2013). Improved conspicuity and delineation of high-grade primary and metastatic brain tumors using "restriction spectrum imaging": quantitative comparison with high B-value DWI and ADC. *AJNR Am J Neuroradiol, 34*(5), 958-964, S951. doi:10.3174/ajnr.A3327

Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., & Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *Neuroimage, 88*, 79-90. doi:10.1016/j.neuroimage.2013.11.027

Zhang, H., Schneider, T., Wheeler-Kingshott, C. A., & Alexander, D. C. (2012). NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *Neuroimage, 61*(4), 1000-1016. doi:10.1016/j.neuroimage.2012.03.072

Table 1. Sample Characteristics

	Discovery Cohort		Replication Cohort		Discovery	Replication
	Male	Female	Male	Female	M vs. F	M vs. F
Sample Size	1788	1611	1741	1657	-	-
Age (years)	9.92 ± 0.62	9.89 ± 0.62	9.93 ± 0.62	9.91 ± 0.61	0.15	0.46
PDS Category					**< 0.001**	**< 0.001**
Prepubertal	1229	473	1172	482		
Early Pubertal	415	374	416	374		
Mid-Pubertal to Postpubertal	88	711	99	742		
CBCL Internalizing Problems Raw	5.06 ± 5.45	5.13 ± 5.61	5.34 ± 5.77	5.22 ± 5.46	0.74	0.54
CBCL Externalizing Problems Raw	4.99 ± 6.18	3.87 ± 5.02	5.17 ± 6.66	3.66 ± 4.92	**< 0.001**	**< 0.001**
CBCL Internalizing Problems T-Score	49.35 ± 10.74	47.73 ± 10.46	49.89 ± 10.69	47.98 ± 10.41	**< 0.001**	**< 0.001**
CBCL Externalizing Problems T-Score	46.39 ± 10.53	45.36 ± 9.65	46.63 ± 10.70	44.84 ± 9.64	**0.003**	**< 0.001**
Mean Head Motion (mm)	1.33 ± 0.49	1.33 ± 0.53	1.34 ± 0.49	1.33 ± 0.51	0.75	0.68
Intracranial Volume (mm^3)	1.59E+06 ± 1.43E+05	1.45E+06 ± 1.28E+05	1.58E+06 ± 1.39E+05	1.46E+06 ± 1.28E+05	**< 0.001**	**< 0.001**
Race					0.49	0.50
Asian	38	39	38	40		
Black	253	243	253	260		
White	1214	1054	1155	1058		
Other/Mixed	283	275	295	299		
Ethnicity					0.32	0.84
Hispanic or Latino/Latina	342	330	361	339		
Not Hispanic or Latino/Latina	1446	1281	1380	1318		
Household Income					0.77	0.50
Less than $50k	512	477	518	503		
Between $50k - $100k	508	443	497	495		
More than $100k	768	691	726	659		
Parental Education					0.36	0.74
Less than HS Diploma	57	71	55	60		
HS Diploma / GED	140	131	160	133		
Some College	489	414	454	430		
Bachelor Degree	467	420	463	443		
Post-Graduate Degree	635	575	609	591		

Descriptive statistics are presented as mean ± standard deviation for continuous variables and as subject counts for discrete variables. PDS = Pubertal Development Scale. CBCL = Child Behavior Checklist. HS = High School. GED = General Education Development test. PDS data was missing for 109 subjects in the discovery cohort and 113 subjects in the replication cohort. CBCL data was missing for 1 subject in the replication cohort

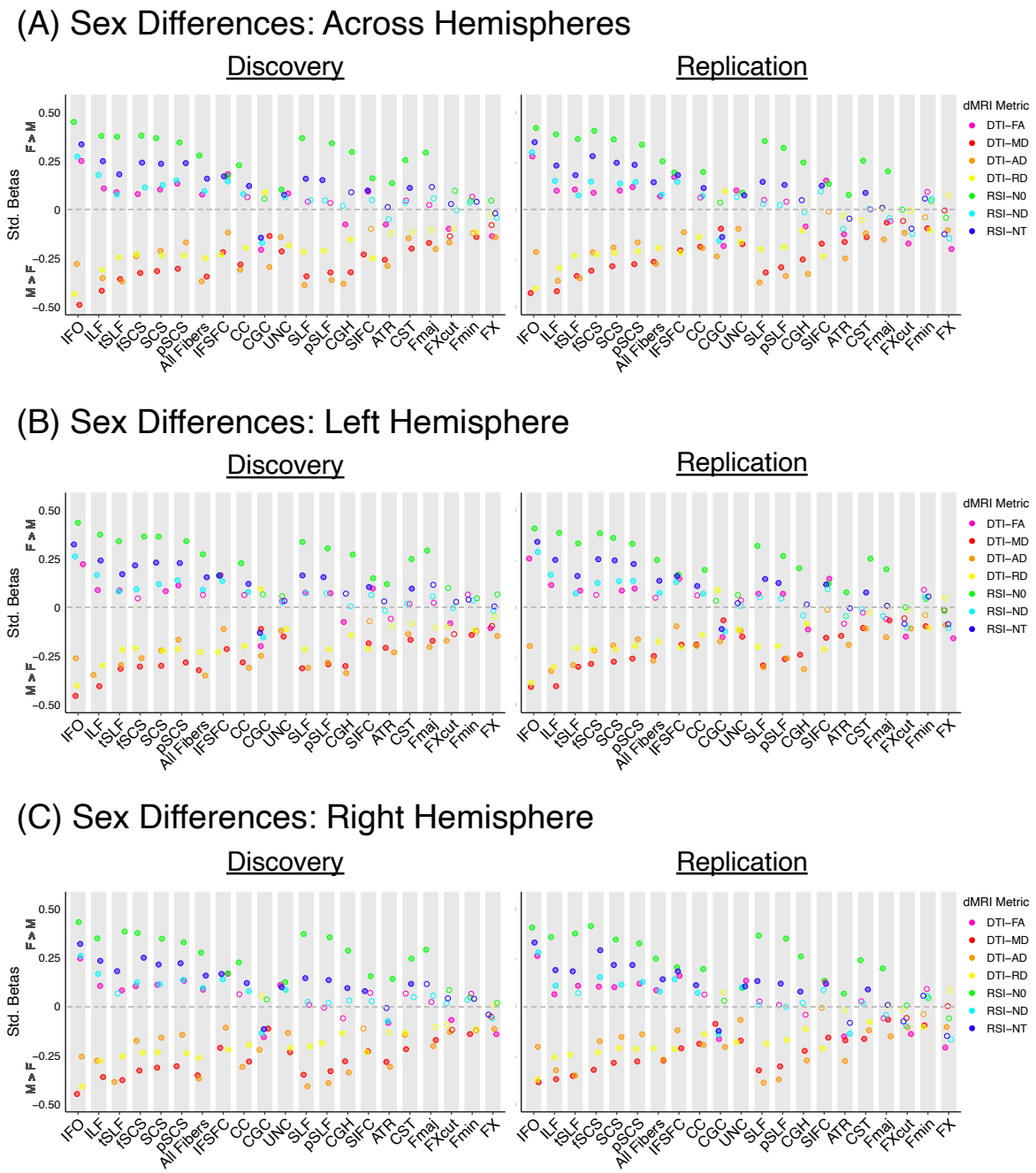

Figure 1. Sex differences in white matter microstructure. The effect size, significance, and replicability of sex differences in DTI and RSI metrics are depicted separately for the discovery cohort (left) and replication cohort (right) for (A) bilateral white matter tract ROIs, (B) left hemisphere white matter tract ROIs, and (C) right hemisphere white matter tract ROIs; commissural tracts are included in the left and right hemisphere graphs for completeness only. Positive standardized betas indicate Female > Male, and negative standardized betas indicate Male > Female. DTI metrics are depicted in warm colors, and RSI metrics in cool colors. Filled circles indicate the association was both significant in the discovery cohort after FDR correction across the number of ROIs (q<0.05) and demonstrated p<0.05 in the replication cohort. *Diffusion-weighted MRI abbreviations*: ROI = region of interest, dMRI = diffusion-weighted MRI, DTI = diffusion tensor imaging, FA = fractional anisotropy, MD = mean diffusivity, AD = axial diffusivity, RD = radial diffusivity, RSI = restriction spectrum imaging, N0 = normalized isotropic, ND = normalized directional, NT = normalized total. *White matter tract ROI abbreviations*: ATR = anterior thalamic radiation, CC = corpus callosum, CGC = cingulum (cingulate), CGH = cingulum (parahippocampal), CST = corticospinal/pyramidal tract, Fmaj = *forceps major*, Fmin = *forceps minor*, fSCS = superior corticostriate (frontal cortex), FX = fornix, FXcut = fornix (excluding fimbria), IFO = inferior fronto-occipital fasciculus, IFSFC = inferior frontal superior frontal cortex, ILF = inferior longitudinal fasciculus, pSCS = superior corticostriate (parietal cortex), pSLF = superior longitudinal fasciculus (parietal), SCS = superior corticostriate, SIFC = striatal inferior frontal cortex, SLF = superior longitudinal fasciculus, tSLF = superior longitudinal fasciculus (temporal), UNC = uncinate fasciculus.

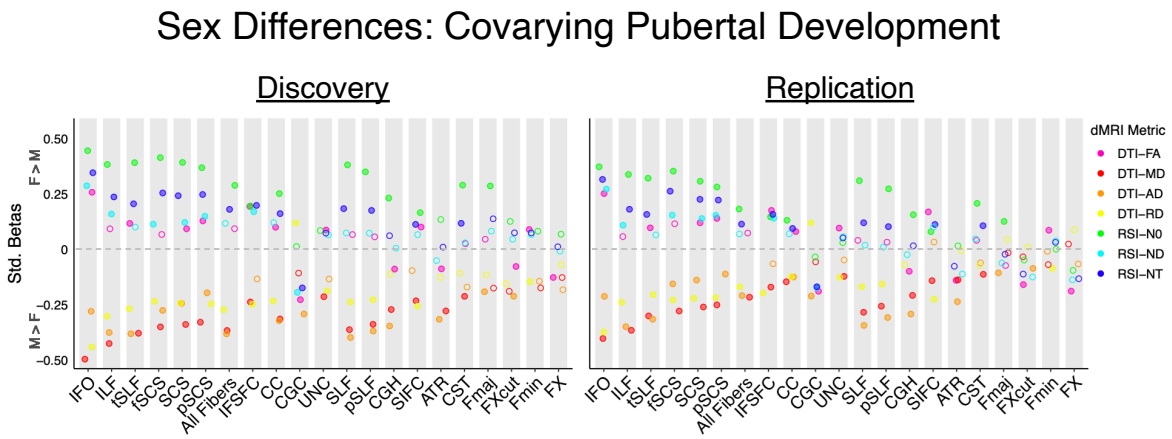

Figure 2. Sex differences in white matter microstructure, controlling for pubertal development. The effect size, significance, and replicability of sex differences in DTI and RSI metrics are depicted separately for the discovery cohort (left) and replication cohort (right) for bilateral white matter tract ROIs. Positive standardized betas indicate Female > Male, and negative standardized betas indicate Male > Female. DTI metrics are depicted in warm colors, and RSI metrics in cool colors. Filled circles indicate the association was both significant in the discovery cohort after FDR correction across the number of ROIs (q<0.05) and demonstrated p<0.05 in the replication cohort. *Diffusion-weighted MRI abbreviations*: ROI = region of interest, dMRI = diffusion-weighted MRI, DTI = diffusion tensor imaging, FA = fractional anisotropy, MD = mean diffusivity, AD = axial diffusivity, RD = radial diffusivity, RSI = restriction spectrum imaging, N0 = normalized isotropic, ND = normalized directional, NT = normalized total. *White matter tract ROI abbreviations*: ATR = anterior thalamic radiation, CC = corpus callosum, CGC = cingulum (cingulate), CGH = cingulum (parahippocampal), CST = corticospinal/pyramidal tract, Fmaj = *forceps major*, Fmin = *forceps minor*, fSCS = superior corticostriate (frontal cortex), FX = fornix, FXcut = fornix (excluding fimbria), IFO = inferior fronto-occipital fasciculus, IFSFC = inferior frontal superior frontal cortex, ILF = inferior longitudinal fasciculus, pSCS = superior corticostriate (parietal cortex), pSLF = superior longitudinal fasciculus (parietal), SCS = superior corticostriate, SIFC = striatal inferior frontal cortex, SLF = superior longitudinal fasciculus, tSLF = superior longitudinal fasciculus (temporal), UNC = uncinate fasciculus.

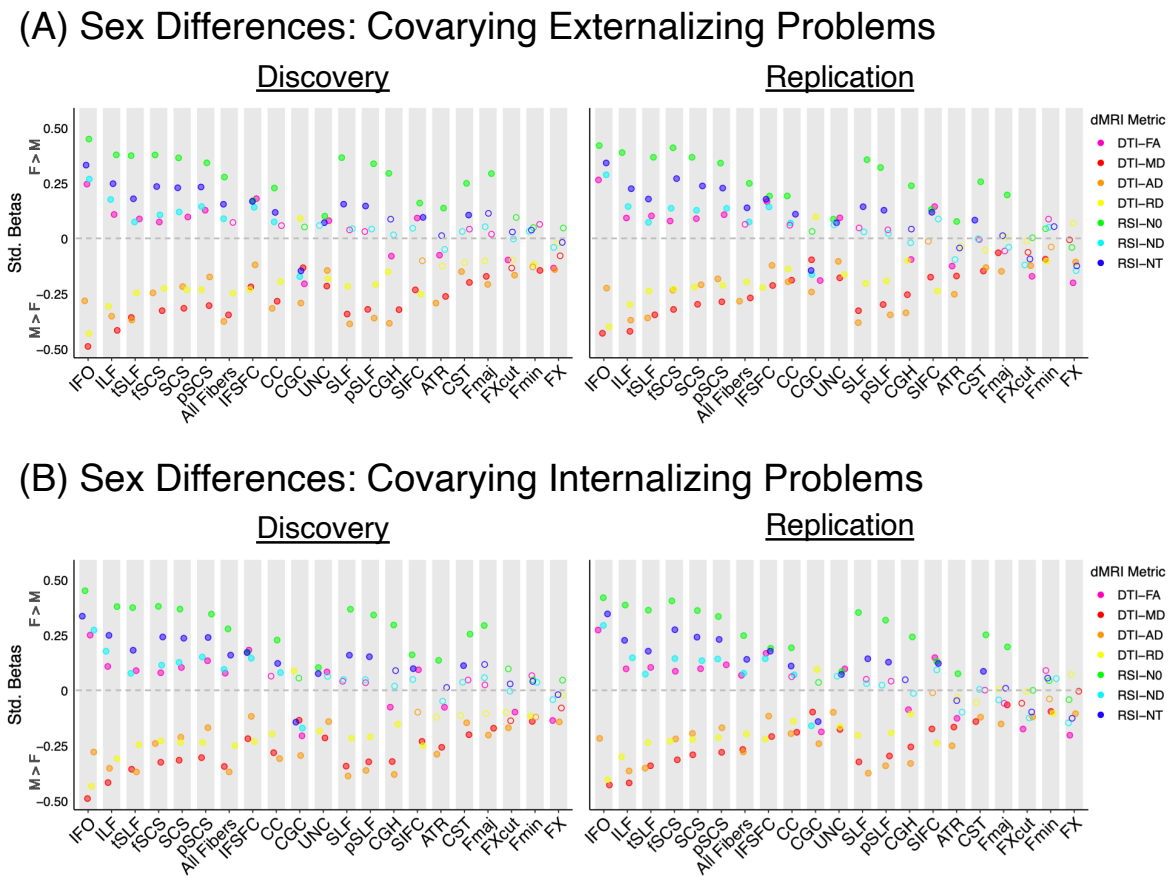

Figure 3. Sex differences in white matter microstructure, controlling for dimensional externalizing and internalizing problems. The effect size, significance, and replicability of sex differences in DTI and RSI metrics are depicted separately for the discovery cohort (left) and replication cohort (right) for bilateral white matter tract ROIs when controlling for (A) dimensional externalizing problems and (B) dimensional internalizing problems. Positive standardized betas indicate Female > Male, and negative standardized betas indicate Male > Female. DTI metrics are depicted in warm colors, and RSI metrics in cool colors. Filled circles indicate the association was both significant in the discovery cohort after FDR correction across the number of ROIs (q<0.05) and demonstrated p<0.05 in the replication cohort. *Diffusion-weighted MRI abbreviations*: ROI = region of interest, dMRI = diffusion-weighted MRI, DTI = diffusion tensor imaging, FA = fractional anisotropy, MD = mean diffusivity, AD = axial diffusivity, RD = radial diffusivity, RSI = restriction spectrum imaging, N0 = normalized isotropic, ND = normalized directional, NT = normalized total. *White matter tract ROI abbreviations*: ATR = anterior thalamic radiation, CC = corpus callosum, CGC = cingulum (cingulate), CGH = cingulum (parahippocampal), CST = corticospinal/pyramidal tract, Fmaj = *forceps major*, Fmin = *forceps minor*, fSCS = superior corticostriate (frontal cortex), FX = fornix, FXcut = fornix (excluding fimbria), IFO = inferior fronto-occipital fasciculus, IFSFC = inferior frontal superior frontal cortex, ILF = inferior longitudinal fasciculus, pSCS = superior corticostriate (parietal cortex), pSLF = superior longitudinal fasciculus (parietal), SCS = superior corticostriate, SIFC = striatal inferior frontal cortex, SLF = superior longitudinal fasciculus, tSLF = superior longitudinal fasciculus (temporal), UNC = uncinate fasciculus.