



Published in final edited form as:

Autism Res. 2017 October ; 10(10): 1606–1620. doi:10.1002/aur.1807.

White matter compromise in autism? Differentiating motion confounds from true differences in diffusion tensor imaging

Ms. Seraphina K. Solders [student], Drs. Ruth A. Carper, and Ralph-Axel Müller

Brain Development Imaging Laboratory, Department of Psychology, San Diego State University, San Diego, CA, USA

Scientific abstract

Background—Common findings from diffusion tensor imaging (DTI) in autism spectrum disorder (ASD) include reduced fractional anisotropy (FA), and increased mean and radial diffusivity (MD, RD) of white matter tracts. However, findings may be confounded by head motion. We examined how group-level motion matching affects DTI comparisons between ASD and typically developing (TD) groups.

Methods—We included 57 ASD and 50 TD participants, comparing 3 subsets at increasing levels of motion-matching stringency: full sample (FS); quality-controlled (QC); and quantitatively-matched (QM). Groups were compared on diffusivity measures using Tract-Based Spatial Statistics (TBSS) and probabilistic tractography. Two methods for estimating diffusivity were compared: *dti-fit* and *restore*.

Results—**TBSS**: In set FS, FA was reduced in the ASD compared to the TD group throughout the right hemisphere. This effect was less extensive in set QC and absent in set QM. However, effect sizes remained stable or increased with better quality-control in some regions. **Tractography**: In set QM, MD was significantly higher in ASD overall and RD was higher in bilateral ILF. Effects were more robust in QM than in FS or QC sets. Effect sizes in several tracts increased with stringent quality matching. *Restore* improved tensor estimates, with some increases in effect sizes, but did not fully compensate for reduced quality.

Conclusions—Findings suggest that some previously reported DTI findings for ASD may have been confounded by motion. However, effects in tightly matched subset indicate that tract-specific anomalies probably do exist in ASD. Our results highlight the need for careful quality-control and motion-matching.

Keywords

diffusion; autism; motion; MRI; white matter; artifact

Correspondence: Ruth A. Carper, Brain Development Imaging Laboratory, Department of Psychology, San Diego State University, 6363 Alvarado Ct., Suite 200, San Diego, CA 92120; rcarper@mail.sdsu.edu.

Financial Disclosures: Drs. Carper, Müller, and Miss Solders report no biomedical financial interests or potential conflicts of interest.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder defined by impaired social communication and interaction as well as restricted and repetitive behaviors and interests (American Psychological Association, 2010). There is ample neuroimaging evidence of atypical brain network connectivity in ASD from functional magnetic resonance imaging (fMRI) and diffusion tensor imaging (DTI). Volumetric MRI and head circumference studies of ASD have consistently indicated a period of accelerated brain growth during infancy and early childhood, followed by abnormally slow growth in later years (Carper et al., 2002; Sparks et al., 2002; Courchesne et al., 2003; Schumann et al., 2010; Hazlett et al., 2011; Hazlett et al., 2012; Shen et al., 2013). Similar patterns have been suggested in structural connectivity studies using DTI to characterize the axonal organization and microstructural properties of white matter. DTI studies of ASD have reported increased fractional anisotropy (FA), potentially an index of premature white matter development, in infants and toddlers (Weinstein et al., 2011; Wolff et al., 2012; Solso et al., 2014), but reduced FA in several white matter pathways later in life (Sundaram et al., 2008; Noriuchi et al., 2010; Shukla et al., 2011; Kleinhans et al., 2012). Common white matter findings in children and adolescents with ASD also include increased mean diffusivity (MD) and radial diffusivity (RD) (Lee et al., 2009; Shukla et al., 2010; Groen et al., 2011; Jou et al., 2011). This pattern of findings has been reported using a variety of analysis approaches.

However, as DTI measures the random motion (diffusion) of protons, with scan times usually around 7-12 minutes, the technique is particularly sensitive to head motion. Depending on the angular resolution chosen, a sequence includes anywhere from 15 to over 100 whole head acquisitions, each sensitive to a different direction of water diffusion. Head motion during the DTI sequence can result in misalignment between these acquisitions and loss of signal intensity in individual slices or portions of slices (slice 'dropout'), resulting in false estimation of diffusion measures (Tournier et al., 2011; Ling et al., 2012; Yendiki et al., 2013; Koldewyn et al., 2014). This is especially problematic in the study of clinical populations like ASD, who may tend to move more than healthy control participants, and in the study of children in general. In these cases, the quality of DTI data may differ systematically between groups, and detected group differences may reflect differences in data quality rather than underlying neurobiology.

While awareness of motion confounds has increased recently in the fMRI connectivity literature, following a series of methodological studies (Power et al., 2015), few DTI studies to date have considered the effects of motion and even fewer have implemented motion measures statistically. Of the 48 DTI studies of ASD included in the exhaustive review by Travers et al. (2012), only a handful explicitly evaluated head motion measures or compared motion between groups. Despite growing awareness of motion confounds, only 28 of the 82 additional diffusion papers on ASD published between January 2012 and October 2016 quantitatively compared head motion between groups (Supplementary Table 1).

Ling et al. (2012) found that increased head motion was associated with a positive bias for FA and MD. Conversely, Yendiki et al. (2013) reported a negative bias for FA and little effect on MD. They further demonstrated that between-group differences in average motion

can result in spurious group differences in diffusion measures even between otherwise well-matched typically developing (TD) samples, and that registering diffusion-weighted images to a baseline image does not eliminate this problem. Additionally, Koldewyn et al. (2014) found that, after quantitative motion matching between ASD and TD groups, the common finding of reduced FA in ASD could no longer be replicated in any white matter tracts other than right inferior longitudinal fasciculus. While that finding might be taken as an indicator of near absence of white matter differences in ASD, it is important to note that quality control and motion matching in Koldewyn et al. (2014) resulted in heavy loss in sample size, leaving it open whether null findings may have been related to weakened statistical power.

The present study used data from 107 ASD and typically developing (TD) children and adolescents, analyzed along two pipelines: Tract-based spatial statistics (TBSS) and probabilistic tractography. Specifically, we tested whether group differences in anisotropy and diffusivity measures would be affected by a series of strict quality control measures and group-matching for motion. We assessed both statistical significance (affected by sample sizes) and effect-sizes of group differences (independent of sample size) in our approach.

Methods and Materials

Participants

Scanning was attempted on 77 ASD and 60 TD participants. Twenty-four (18 ASD, 6 TD) did not complete the full diffusion sequence, and technical error eliminated one additional TD scan. Diffusion weighted MRI data were successfully collected from the other 59 ASD and 53 TD participants, ages 7-17 years. Five of these participants were excluded due to: atypical findings on MRI (2 ASD, 2 TD), or having an adopted sibling with ASD (1 TD). The final sample consisted of 57 ASD and 50 TD participants.

All ASD participants met the DSM-5 (American Psychological Association, 2010) criteria for autism spectrum disorder, based on the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2002) and Autism Diagnostic Interview-Revised (ADI-R; Rutter et al., 1995). A clinical psychologist experienced in the diagnosis of ASD confirmed the final diagnosis. Participants with ASD had no other known neurological or psychiatric disorders, and no TD participants had a personal or family history of neuropsychiatric conditions. The Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was used to obtain IQ scores. This study was approved by the University of California, San Diego, and San Diego State University Institutional Review Boards, with written informed consent and assent from all participants and caregivers.

MRI data acquisition

MRI was performed on a GE Discovery MR750 3.0T scanner with an 8-channel head coil. Diffusion weighted images were collected using an echo planar imaging (EPI) pulse sequence with full head coverage, encoded for 61 non-collinear diffusion directions at $b=1000 \text{ s/mm}^2$, and one at $b=0 \text{ s/mm}^2$ (2D EPI; TR=8500ms; TE=84.9ms; flip angle=90°; NEX=1; FOV=24cm; resolution= $1.875 \times 1.875 \times 2 \text{ mm}^3$). A field map was collected with

the same spatial parameters to correct for field inhomogeneities (2D GRASS; TR=1097ms; TE=7.5ms, 9.5ms; flip angle=45°; 2 averages).

Image Analysis

DTI processing was performed using the FMRIB Software Library (FSL, v.5.0; Smith et al., 2004) and AFNI (v.2011_12_21_1014; Cox, 1996). Preprocessing included field-map correction, resampling to a $1 \times 1 \times 1 \text{mm}^3$ resolution, removing skull and other non-brain tissue (FSL BET), and eddy current correction (FSL eddy_correct). FSL's *dti-fit* calculated the diffusion tensor at each voxel and generated maps of axial and radial diffusivity (AD and RD), mean diffusivity (MD), and fractional anisotropy (FA). The FA map was non-linearly registered (FSL FNIRT) to the standard MNI template (FMRIB58-FA), and the resulting transformation matrices were inverted and saved to facilitate fiber tracking in native space using standard-space atlases.

We also assessed the relative effectiveness of a tensor calculation algorithm frequently used for quality improvement. Robust Estimation of Tensors by Outlier Rejection (*restore*) is performed on a voxel-wise basis to identify and exclude outliers from the multiple diffusion directions collected, and calculates the diffusion tensor from the remaining data (Chang et al., 2005). We calculated diffusivity maps using *restore* after the eddy correction step, offering an alternative to *dti-fit*.

Quality Control and Motion Matching

To examine the effects of motion differences between groups, we examined different subsets of the sample matched at increasing levels of stringency based on two forms of quality control. First, quality control was implemented consisting of visual inspection for slice-wise signal dropout, image noise, and shifts of head placement between diffusion volumes. Scans were excluded if they had single slices of signal dropout affecting 10 diffusion directions, multiple slices of signal dropout on 5 diffusion directions, color banding evident in the RGB display of the primary eigenvector image, visible nods or head shakes between diffusion directions, or visible image noise. Second, for greater stringency, quantitative assessment was used in order to match groups on four motion measures (average inter-volume translation, average rotation, proportion of slices affected by signal dropout, severity of signal dropout) as described by Yendiki et al. (2013), with reference to Benner et al. (2011). Group means were matched to less than five percent difference on each motion measure. We examined group differences in the full sample (FS), the quality-controlled (QC) subset of the FS, and the quantitatively-matched (QM) subset of the QC set (Table). At each stage, groups were compared on FA, MD, RD, and AD (from both *dti-fit* and *restore*) using two approaches provided in FSL (Smith et al., 2004): Tract-Based Spatial Statistics (TBSS) and probabilistic tractography (ProbtrackX).

TBSS Analysis

Voxelwise statistical analysis was performed using TBSS (Smith et al., 2006). FA data from each participant were aligned into a common space (FMRIB58-FA) using nonlinear registration. A mean FA image was created and thinned to create a mean FA skeleton representing the centers of tracts common to the sample. Each participant's aligned FA data

were then projected onto this skeleton and resulting data fed into FSL's randomise tool (<http://www.fmrib.ox.ac.uk/fsl/randomise/index.html>; 500 permutations) to perform voxelwise cross-subject statistics. Following the FA analyses, the nonlinear warps and skeleton projections were applied to MD, RD, and AD, and voxelwise statistics were run for each measure with Threshold-Free Cluster Enhancement correction, which fully corrects for multiple comparisons across space. The locations of significant clusters were determined using the Johns Hopkins University White-Matter Tractography Atlas (Hua et al., 2008; Mori et al., 2008).

While increasing stringency of motion matching reduced artifactual variance and a likely source of group bias, it also reduced sample sizes and statistical power. We therefore

calculated voxel-wise effect size maps (voxel-wise Cohen's d : $\frac{\bar{x}_{ASD} - \bar{x}_{TD}}{\text{pooled } s.d.}$ at each stage of motion matching for sample-size independent comparisons. Significant clusters from the QC subset were then used as selection masks to determine the average effect sizes of these regions across the FS, QC, and QM subsets.

Probabilistic Tractography Analysis

Tractography was performed using probabilistic algorithms in FSL (BEDPOSTX, ProbtrackX; Smith et al., 2004). Tracts of interest included the forceps major (Fmaj) and minor (Fmin), the superior longitudinal fasciculus (SLF), the inferior longitudinal fasciculus (ILF), and the uncinate fasciculus (UNC), separately by hemisphere (Figure 1) in order to assess the major association pathways. Fasciculi were identified using seed and target masks defined in standard MNI space in accordance with Wakana et al. (2007). Masks were transformed to individual subject space to conduct tractography.

For each tract and participant, 3000 streamlines were initiated per seed voxel, (step-length=0.5, curvature threshold=0.2), connectivity distribution was corrected for the length of the pathway, and tracking was constrained by anisotropy values. For all tracts except the Fmaj and Fmin, a midsagittal mask was used to exclude any samples that crossed cerebral hemispheres. Resulting probability maps were transformed back into standard MNI space for all subsequent steps.

To remove extraneous fibers, a threshold of 0.01% of initiated streamlines was applied to each participant's connectivity maps of all tracts except the SLF, which required a higher threshold (0.1%) for anatomical accuracy. The resulting connectivity maps were binarized and applied to the corresponding DTI maps to obtain mean FA, MD, AD, and RD values within each tract on a per-subject basis. Finally, these values were used to calculate Cohen's d for each tract at each matching stage for sample-size-independent comparisons.

Results

Quality Control and Motion Matching

The FS, without exclusion based on quality control, showed substantial group differences (ASD>TD; 7-39%) on the four motion measures (Table, panel A), with significant differences for the two measures of signal dropout. Careful quality control (QC subset)

greatly reduced the inclusion of outliers, as seen in reduced ranges of motion measures, but groups still differed significantly and substantially (10-38%) on two motion measures (Table, panel B). Stringent quantitative matching (QM subset) reduced group differences to less than 5% on all four motion measures (all $p > .2$; Table, panel C).

We compared retained and excluded participants to determine if they differed. Comparing the QC subset to those excluded during visual quality control, groups did not differ significantly on age, IQ, or ASD symptomatology (on ADOS Communication and Social Interaction, ADOS Restricted and Repetitive Behaviors, and SRS Total scores; Supplementary Table 2). Further, the final subset (QM) was compared to all participants excluded during either matching step. Again, there were no significant differences in the ASD or the TD group (Supplementary Table 3).

TBSS Analysis

Tensor estimation with dti-fit—In the FS sample, FA was significantly reduced in the ASD compared to the TD group in several clusters (total volume 7089 mm³), almost exclusively in the right hemisphere (Figure 2). In the QC sample, FA was still significantly lower in ASD in multiple clusters, but areas of difference were less extensive (total volume 4077 mm³). Finally, in the QM subset, no significant differences were found for FA. Regardless of the stringency of quality control (FS, QC, or QM), none of the comparisons resulted in significant group differences for MD, AD, or RD.

Progressively smaller sample-sizes from FS to QC and QM sets negatively affected statistical power. For sample-size-independent comparisons, we therefore calculated voxelwise effect sizes of group differences for each set. While statistical significance generally decreased across sets, effect sizes increased in many regions. For illustration, clusters that showed significant FA differences in the QC set were used as selection masks, applied to the effect size maps for each set, and averaged within clusters (Figure 3a). Notably, moderate to large effect sizes (Cohen's d 0.4-0.8) were seen for FA in the QM set in five out of six clusters even though differences did not reach significance. Effect sizes were also comparable or greater for the QM than for the FS set in five out of six clusters, suggesting that group differences were not motion artifacts. The same selection masks were used to summarize effect sizes for the other dependent measures (MD, AD, RD) even though these did not reach statistical significance. Strikingly, effect sizes for RD *increased* with improved motion control, reaching medium ($>.4$) effect sizes in the QM sample in 5 out of 6 clusters.

Tensor estimation with restore—TBSS was run independently using diffusivity estimates from *restore*. No significant group effects were found in any of the three datasets for any dependent measure. When thresholds were lowered to $p = .1$, the spatial distribution of group differences on FA was seen to be similar to that for *dti-fit* estimates in both the FS and QC sets, though not reaching significance.

Effect sizes were calculated and the same selection masks were used for comparison purposes. Effect sizes were quite similar to those calculated from *dti-fit* estimates, and showed similar patterns across datasets (Figure 3b).

Probabilistic Tractography Analysis

Tensor estimation with *dti-fit*—ANOVAs were performed on results from probabilistic tractography, with separate tests for each dependent variable (FA, MD, AD, RD). In 4 analyses, cerebral hemisphere (left, right) and tract (SLF, ILF, UNC) were the within-subjects variables, and group (ASD, TD) was the between-subjects variable. The inter-hemispheric tracts (F_{maj} , F_{min}) were tested with tract (F_{maj} , F_{min}) as the only within-subjects variable. All analyses were performed at each stage of quality control (FS, QC, QM). The results of all ANOVAs are summarized in Supplementary Table 4.

In the most stringently matched QM set, analysis of FA in the intra-hemispheric tracts showed no significant effects related to group. In contrast, MD was significantly higher in the ASD than the TD group ($F(1, 60)=4.55, p=.037$). A significant group-by-tract interaction was found for RD in the QM set ($F(2, 59)=3.35, p=.042$), where post-hoc ANOVAs performed separately for each tract revealed increased RD bilaterally in the ASD group compared to the TD group in the ILF only ($p=.008$). No significant effects related to group were found for AD.

Notably, group and group-by-tract effects from the corresponding ANOVAs for FS and QC sets were less robust than for QM on MD, RD, and AD, despite the larger sample sizes. This pattern was further supported by effect sizes, plotted for individual tracts in Figure 4a. In most tracts, MD, AD and RD effect sizes were greatest for the QM set, with left and right ILF showing a particularly striking increase in effect size linked to improvements in motion matching. In contrast, effect sizes for FA measures tended to diminish with improved motion matching.

Finally, group effects for inter-hemispheric tracts approached significance in the QM set for both MD ($F(1, 62)=3.94, p=.052, ASD>TD$) and RD ($F(1, 62)=3.86, p=.054, ASD>TD$). Group-related effects did not approach significance for either FA or AD. Once again, effects from the corresponding ANOVAs for the less stringently matched FS and QC sets were less robust. Examination of effect sizes in the different samples showed the largest effect sizes were most often in the stringently matched QM set (Figure 4a).

Tensor estimation with *restore*—The group effects seen in the QM set with *dti-fit* were not detected with *restore* (Supplementary Table 5). For intra-hemispheric tracts, the main effect of group reached $p=.059$, for MD and $p=.082$ for RD. No other group effects or interactions approached significance. However, in the larger QC set, a significant group by tract interaction was found for FA ($p=.037$), revealing a significant group effect for the ILF ($ASD<TD$) in post-hoc ANOVAs. A similar interaction for RD ($p=.03, ASD>TD$) also reflected an effect in ILF. Group and group-by-tract effects for the FS set were less robust for MD, AD, and RD. For inter-hemispheric tracts, group effects approached significance in the QM set for both MD ($p=.063$) and RD ($p=.09$). In the QC set, RD reached significance ($F(1, 73)=4.31, p=.041$), with a marginal effect for MD ($F(1, 73)=2.71, p=.10$). Differences in the FS set were again less robust, and group-related effects did not approach significance for either FA or AD in any set. For both inter- and intra-hemispheric tracts, effect sizes (Figure 4b) again tended to be greatest in the QM set for MD, AD, and RD although the

pattern was not as consistent as with *dti-fit* estimates. FA effect sizes again tended to diminish with improved motion matching.

Discussion

We examined the effects of quality control and stringent motion matching in comparisons between ASD and TD groups. The main findings can be summarized as follows: First, we found evidence of false-positive effects when groups were not stringently matched for motion, thus replicating aspects of a previous report by Koldewyn et al. (2014). Significant between-group differences were found for FA without motion matching, but significance was absent when stringent matching was applied. Strikingly, this suggests that many previously reported DTI findings, from ASD studies that did not apply stringent motion matching, may be artifactual. Second, however, our findings showed that effect sizes in several clusters (from TBSS) and tracts (from tractography) were remarkably stable or even increased with stringency of quality control for several DTI indices (Figures 1 and 2). This suggests that white matter compromise in ASD exists, but requires careful quality control for reliable detection (i.e. risk of false-negative effects).

Many group effects disappear with motion matching

Although exact findings from TBSS and Protrack varied, both methods yielded group effects that reached statistical significance in the FS set, but ‘disappeared’ below significance thresholds with tighter motion matching. In both methods, the common finding of reduced FA in ASD was replicated in the FS – quite extensively with TBSS (Supplementary Figure 2), more restricted with tractography, detected only in right ILF— but these significant effects were not seen in the QM set for either method.

Notably, no significant effects for FA in the QM set were seen in some regions for which reduced FA had been reported in many previous studies, before recent motion matching procedures became available. These regions include corpus callosum, cingulum bundle, and SLF (see review by Travers et al., 2012). Although our null findings in the QM set do not demonstrate complete absence of white matter compromise in these tracts (as discussed below), they indicate that highly significant effects reported in some studies with smaller samples (Cheon et al., 2011; Poustka et al., 2012; Cascio et al., 2013; Jou et al., 2016) may have been inflated by group differences in head motion.

Effect sizes increase with motion matching in some regions

While motion control eliminated some significant effects, it is important to note that it also revealed effects not seen in less tightly matched data. With tractography, no significant group effects were detected in the FS set for either MD or RD; however, in the QM set MD was significantly increased in ASD. Increased RD was nearly significant ($p=.052$) for forceps major and minor, and a significant group-by-tract interaction on RD was revealed for intrahemispheric tracts in QC and QM (but not FS) sets. Effect sizes showed a corresponding pattern, increasing with improved motion matching for multiple measures and in multiple tracts and clusters in both tractography and TBSS analyses (Figures 1-2).

Thus, while some of the effects commonly reported in the literature may have been confounded by head motion, our findings do not support absence of abnormal structural connectivity in ASD. Rather, they suggest true differences in ILF and other tracts, and broadly higher MD in ASD (discussed below). The lack of additional statistically significant group findings in the presence of increasing effect sizes, particularly for RD and AD, can be attributed to reduced sample size (c. 40% lower in QM than FS set) and may reflect type 2 error (false negatives). The lack of additional significant findings in the study by Koldewyn et al. (2014) may be similarly attributed to reduced statistical power. Their sample decreased by c. 80% with motion matching, the best-matched set including only 38 participants.

Algorithms for diffusion tensor estimation

With in-scanner motion leading to changes in diffusivity measures, techniques for improving data quality post-hoc can be beneficial. The commonly used technique *restore* reduces the impact of motion artifacts on a volume-wise (diffusion direction) and voxel-wise basis. If *restore* fully compensated for group differences in motion, we would expect constant effect sizes across our 3 quality-based datasets. The QC and QM sets appeared more similar with *restore* than *dti-fit*, especially on RD, but still differed on several tracts. This suggests that *restore* improves reliability of tensor measures, but does not fully compensate for reduced data quality. The use of *restore* is therefore commendable, but only when applied to motion-matched and quality-screened groups.

Participants with greater motion did not differ on demographics or symptom severity

Remarkably, subsamples with higher vs. lower motion (those excluded vs. those included in the QM set) did not differ significantly on age, IQ, or symptom severity. This indicates that careful motion matching may *not* introduce a sampling bias, at least among high-functioning children as examined here. This may be attributed, in part, to our use of mock-scanner training whereby all participants underwent at least one practice session in a simulated scanner environment, with feedback –continued viewing of a video– for maintaining constant head position.

Sources of motion and reducing motion

While we assumed subject-initiated motion to be the primary source of in-scanner motion, mechanical causes should also be considered. The rapid gradient switching of diffusion sequences is notorious for inducing scanner vibration. When the gradient is applied maximally to a single gradient coil, or high b-values are used, localized areas of signal dropout can occur from vibrationally-induced shear-waves in brain tissue (Gallichan et al., 2010; Berl et al., 2015). This is not problematic under normal circumstances and typical b-values, but may be an issue with mechanical deficiencies (e.g. a bed-locking problem on some Siemens systems; Gallichan et al., 2010), or pediatric populations (Berl et al., 2015). Although rare, these issues further underscore the need for careful group matching – including quantification of signal dropouts– to balance outlier signals regardless of source.

When planning diffusion studies, sequences should be selected to minimize and recover from all sources of motion artifacts. Gallichan et al. (2010) recommend acquiring full k-space data and Berl et al. (2015) found that higher spatial resolution sequences may be less

susceptible. Longer echo times can reduce vibrational artifacts, but with a trade off in signal-to-noise. Sampling from a higher number of directions can help with intermittent subject-induced artifacts as this essentially increases the number of images available for averaging, or multiple acquisitions of a single sequence may be collected and averaged. Such improvements should be considered, if compatible with limited duration of MRI sessions in pediatric and other challenging populations.

Subject-initiated motion can be minimized mechanically with padding, bite-bars and other methods. We used mock-scanner training to help reduce motion artifacts in our pediatric and ASD populations (as well as padding during imaging). Some research groups use anesthesia in patient populations, but anesthesia in healthy control groups is prohibited. While this may reduce subject-induced motion in the one group, it may nonetheless result in mismatched groups.

Implications for ASD

The results from the most stringent QM set, expected to represent anatomical differences in ASD with lowest probability of type 1 error, indicate atypically increased MD across multiple tracts and increased RD in bilateral ILF in the ASD group. This is partially consistent with the finding by Koldewyn et al. (2014) which showed reduced FA in the right ILF after quantitative group matching on motion. The ILF is a major pathway of the ventral visual stream, with feedforward and feedback connections between occipital and temporal cortices, and between visual cortices and amygdala (Schmahmann and Pandya, 2006). The tract plays a key role in object and face recognition (Haxby et al., 2000; Haxby et al., 2001; Hodgetts et al., 2015; Unger et al., 2016), found to be impaired in ASD, especially in regard to facial identities and expressions (Dawson et al., 2002; Dawson et al., 2004; Lerner et al., 2013).

In addition to significant findings in bilateral ILF, effect sizes were medium to large in the QM set in tractography or TBSS analyses (or both) for right UNC (MD, RD), forceps major (FA, MD, RD), right IFOF (RD), and right SLF (FA, RD). The UNC connects orbital and medial frontal lobe, involved in emotional valuation, reward circuits, and self-regulation, with rostral temporal areas, involved in sound recognition and object identification, and amygdala (Quirk and Beer, 2006; Schmahmann and Pandya, 2006; Thiebaut de Schotten et al., 2012). The UNC thus plays a key role in attaching emotional relevance to visual and auditory stimuli (Barbas and De Olmos, 1990; Ghashghaei and Barbas, 2002; Schmahmann and Pandya, 2006; Von Der Heide et al., 2013; Oishi et al., 2015), which may be impaired in ASD (Green et al., 2013; Blasi et al., 2015; Kana et al., 2016). Diffusion and dissection studies indicate that the IFOF connects some of these same inferior frontal regions with ventral occipital areas (Forkel et al., 2014), implicating relevance in interpretation of facial affect and emotional valuation of visual stimuli (Philippi et al., 2009; Unger et al., 2016). While the left SLF has a prominent role in language function (Catani et al., 2005; Breier et al., 2008; Caverzasi et al., 2016), the larger group differences ($d > .4$) occurred in the right hemisphere where SLF may primarily support spatial attention (Gitelman et al., 1999; Mapstone et al., 2003; Lunven and Bartolomeo, 2016) and aspects of communication such as prosody, gesture, and facial expression (Ahern et al., 1991; Dara et al., 2014; Sammler et

al., 2015), which may be impaired in ASD (Mann and Walker, 2003; Grossman et al., 2013; Robertson et al., 2013; Oerlemans et al., 2014). Finally, the forceps major provides interhemispheric integration between left and right visual fields (Aboitiz et al., 1992; Gazzaniga, 1995; Arguin et al., 2000; Schmahmann and Pandya, 2006; Yamada et al., 2015), and any compromise may relate to atypical visual spatial attention and extensive differences in functional asymmetries in ASD (Haist et al., 2005; Cardinale et al., 2013).

Limitations

Our ASD sample was mostly limited to high-functioning individuals who could hold (mostly) still during scanning. Therefore, our findings may not generalize to lower-functioning children with ASD, who may exhibit different or more pronounced white matter abnormalities. Strict quality control and motion matching for the QM set came at the expense of reduced sample size and power. While this reduced the risk of false positive findings, it was also likely to increase the risk of false negatives. As suggested by effect sizes, some true group differences may have therefore failed to reach significance.

Diffusion imaging in ASD: Challenges and Perspectives

Our findings show that group differences in head motion may result in artifactual or inflated DTI findings. This is consistent with earlier reports (Yendiki et al., 2013; Koldewyn et al., 2014) and may suggest that many oft-cited findings from DTI studies that failed to implement tight group matching may be incorrect. Notably however, our findings go well beyond those in earlier reports (Yendiki et al., 2013; Koldewyn et al., 2014) and show that effects in a number of tracts (beyond right ILF) are stable or become more pronounced with motion matching. The fact that some of these nevertheless failed to reach significance indicates, not only purely statistical limitations (which are of little interest from the neurobiological perspective), but likely also the true and expected heterogeneity of the disorder. If ASD is considered a clinical umbrella term encompassing potentially hundreds of different and rare etiologies (Geschwind and State, 2015), it is expected that any tract-specific white matter anomaly may be detected only in a subset of these. Differential effects from other etiologies comprised under the ASD umbrella will reduce the probability of detecting ‘significant’ differences at the group level. The field thus faces paradoxical challenges: On the one hand, variability and heterogeneity indicate the need for very large samples; on the other, the expectation of many rare variants highlights the need to study children with ASD *as individuals*. While a focus on the individual was beyond the scope of the present investigation on the impact of motion on group comparisons, such added focus in future studies may contribute to a better appreciation of the complexity of the disorder.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors especially thank the children and families who participated in the study. Additional thanks to Dr. Rebecca Theilmann and Dr. Robert Bussell for conversations on MRI physics and potential sources of artifact. The work presented here was supported by National Institutes of Health, grants R01 MH081023 (PI: RAM), K01 MH097972 (PI: Inna Fishman), and R25 GM058906 (in support of SKS). Preliminary findings from this study

appeared in poster form with the same authors at the International Meeting for Autism Research in 2015, and at the Society for Neuroscience Conference in 2016.

Grant sponsor: National Institutes of Health; *Grant number:* R01 MH081023, K01 MH097972 (PI: Inna Fishman), R25 GM058906 (SKS only)

References

- Aboitiz F, Scheibel AB, Fisher RS, Zaidel E. Fiber composition of the human corpus callosum. *Brain research*. 1992; 598:143–153. [PubMed: 1486477]
- Ahern GL, Schomer DL, Kleefield J, Blume H, Cosgrove GR, Weintraub S, et al. Right hemisphere advantage for evaluating emotional facial expressions. *Cortex; a journal devoted to the study of the nervous system and behavior*. 1991; 27:193–202. [PubMed: 1879148]
- American Psychological Association Committee on Legal Issues. *Ethical Principles of Psychologists and Code of Conduct 2010*. American Psychological Association. 2010; 57:1–18.
- Arguin M, Lassonde M, Quattrini A, Del Pesce M, Foschi N, Papo I. Divided visuo-spatial attention systems with total and anterior callosotomy. *Neuropsychologia*. 2000; 38:283–291. [PubMed: 10678694]
- Barbas H, De Olmos J. Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. *The Journal of comparative neurology*. 1990; 300:549–571. [PubMed: 2273093]
- Benner T, van der Kouwe AJW, Sorensen AG. *Diffusion Imaging with Prospective Motion Correction and Reacquisition*. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*. 2011
- Berl MM, Walker L, Modi P, Irfanoglu MO, Sarlls JE, Nayak A, et al. Investigation of vibration-induced artifact in clinical diffusion-weighted imaging of pediatric subjects. *Human brain mapping*. 2015; 36:4745–4757. [PubMed: 26350492]
- Blasi A, Lloyd-Fox S, Sethna V, Brammer MJ, Mercure E, Murray L, et al. Atypical processing of voice sounds in infants at risk for autism spectrum disorder. *Cortex; a journal devoted to the study of the nervous system and behavior*. 2015; 71:122–133. [PubMed: 26200892]
- Breier JI, Hasan KM, Zhang W, Men D, Papanicolaou AC. Language dysfunction after stroke and damage to white matter tracts evaluated using diffusion tensor imaging. *AJNR American journal of neuroradiology*. 2008; 29:483–487. [PubMed: 18039757]
- Cardinale RC, Shih P, Fishman I, Ford LM, Müller RA. Pervasive rightward asymmetry shifts of functional networks in autism spectrum disorder. *JAMA Psychiatry*. 2013; 70:975–982. [PubMed: 23903586]
- Carper RA, Moses P, Tigue ZD, Courchesne E. Cerebral Lobes in Autism: Early Hyperplasia and Abnormal Age Effects. *NeuroImage*. 2002; 16:1038–1051. [PubMed: 12202091]
- Cascio C, Gribbin M, Gouttard S, Smith RG, Jomier M, Field S, et al. Fractional anisotropy distributions in 2- to 6-year-old children with autism. *Journal of intellectual disability research : JIDR*. 2013; 57:1037–1049. [PubMed: 22998325]
- Catani M, Jones DK, ffytche DH. Perisylvian language networks of the human brain. *Annals of neurology*. 2005; 57:8–16. [PubMed: 15597383]
- Caverzasi E, Hervey-Jumper SL, Jordan KM, Lobach IV, Li J, Panara V, et al. Identifying preoperative language tracts and predicting postoperative functional recovery using HARDI q-ball fiber tractography in patients with gliomas. *Journal of neurosurgery*. 2016; 125:33–45. [PubMed: 26654181]
- Chang LC, Jones DK, Pierpaoli C. RESTORE: robust estimation of tensors by outlier rejection. *Magnetic resonance in medicine*. 2005; 53:1088–1095. [PubMed: 15844157]
- Cheon KA, Kim YS, Oh SH, Park SY, Yoon HW, Herrington J, et al. Involvement of the anterior thalamic radiation in boys with high functioning autism spectrum disorders: A Diffusion Tensor Imaging study. *Brain Research*. 2011; 1417:77–86. [PubMed: 21890117]
- Courchesne E, Carper R, Akshoomoff N. Evidence of Brain Overgrowth in the First Year of Life in Autism. *JAMA : the journal of the American Medical Association*. 2003; 290:1–8.

- Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers & Biomedical Research*. 1996; 29:162–173. [PubMed: 8812068]
- Dara C, Bang J, Gottesman RF, Hillis AE. Right hemisphere dysfunction is better predicted by emotional prosody impairments as compared to neglect. *Journal of neurology & translational neuroscience*. 2014; 2:1037. [PubMed: 24945016]
- Dawson G, Carver L, Meltzoff AN, Panagiotides H, McPartland J, Webb SJ. Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Dev*. 2002; 73:700–717. [PubMed: 12038546]
- Dawson G, Webb SJ, Carver L, Panagiotides H, McPartland J. Young children with autism show atypical brain responses to fearful versus neutral facial expressions of emotion. *Developmental Science*. 2004; 7:340–359. [PubMed: 15595374]
- Forkel SJ, Thiebaut de Schotten M, Kawadler JM, Dell'Acqua F, Danek A, Catani M. The anatomy of fronto-occipital connections from early blunt dissections to contemporary tractography. *Cortex*. 2014; 56:73–84. [PubMed: 23137651]
- Gallichan D, Scholz J, Bartsch A, Behrens TE, Robson MD, Miller KL. Addressing a systematic vibration artifact in diffusion-weighted MRI. *Human brain mapping*. 2010; 31:193–202. [PubMed: 19603408]
- Gazzaniga MS. Principles of human brain organization derived from split-brain studies. *Neuron*. 1995; 14:217–228. [PubMed: 7857634]
- Geschwind DH, State MW. Gene hunting in autism spectrum disorder: on the path to precision medicine. *The Lancet Neurology*. 2015; 14:1109–1120. [PubMed: 25891009]
- Ghashghaei HT, Barbas H. Pathways for emotion: Interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience*. 2002; 115:1261–1279. [PubMed: 12453496]
- Gitelman DR, Nobre AC, Parrish TB, LaBar KS, Kim YH, Meyer JR, et al. A large-scale distributed network for covert spatial attention: further anatomical delineation based on stringent behavioural and cognitive controls. *Brain : a journal of neurology*. 1999; 122(Pt 6):1093–1106. [PubMed: 10356062]
- Green SA, Rudie JD, Colich NL, Wood JJ, Shirinyan D, Hernandez L, et al. Overreactive brain responses to sensory stimuli in youth with autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2013; 52:1158–1172. [PubMed: 24157390]
- Groen WB, Buitelaar JK, van der Gaag RJ, Zwiers MP. Pervasive microstructural abnormalities in autism: A DTI study. *Journal of Psychiatry and Neuroscience*. 2011; 36:32–40. [PubMed: 20964953]
- Grossman RB, Edelson LR, Tager-Flusberg H. Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. *Journal of speech, language, and hearing research : JSLHR*. 2013; 56:1035–1044.
- Haist F, Adamo M, Westerfield M, Courchesne E, Townsend J. The functional neuroanatomy of spatial attention in autism spectrum disorder. *Developmental neuropsychology*. 2005; 27:425–458. [PubMed: 15843105]
- Hau J, Sarubbo S, Perchey G, Crivello F, Zago L, Mellet E, et al. Cortical Terminations of the Inferior Fronto-Occipital and Uncinate Fasciculi: Anatomical Stem-Based Virtual Dissection. *Frontiers in Neuroanatomy*. 2016; 10:58. [PubMed: 27252628]
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*. 2001; 293:2425–2430. [PubMed: 11577229]
- Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends in Cognitive Sciences*. 2000
- Hazlett HC, Poe MD, Lightbody AA, Styner M, MacFall JR, Reiss AL, et al. Trajectories of early brain volume development in fragile X syndrome and autism. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2012; 51:921–933. [PubMed: 22917205]
- Hazlett HC, Poe MD, Gerig G, Styner M, Chappell C, Smith RG, et al. Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Archives of general psychiatry*. 2011; 68:467–76. [PubMed: 21536976]

- Von Der Heide RJ, Skipper LM, Klobusicky E, Olson IR. Dissecting the uncinate fasciculus: Disorders, controversies and a hypothesis. *Brain*. 2013
- Hodgetts CJ, Postans M, Shine JP, Jones DK, Lawrence AD, Graham KS. Dissociable roles of the inferior longitudinal fasciculus and fornix in face and place perception. *eLife*. 2015;4.
- Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, et al. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage*. 2008; 39:336–347. [PubMed: 17931890]
- Jou RJ, Mateljevic N, Kaiser MD, Sugrue DR, Volkmar FR, Pelphrey KA. Structural neural phenotype of autism: Preliminary evidence from a diffusion tensor imaging study using tract-based spatial statistics. *American Journal of Neuroradiology*. 2011; 32:1607–1613. [PubMed: 21799040]
- Jou RJ, Reed HE, Kaiser MD, Voos AC, Volkmar FR, Pelphrey KA. White Matter Abnormalities in Autism and Unaffected Siblings. *The Journal of neuropsychiatry and clinical neurosciences*. 2016; 28:49–55. [PubMed: 26404174]
- Kana RK, Patriquin MA, Black BS, Channell MM, Wicker B. Altered Medial Frontal and Superior Temporal Response to Implicit Processing of Emotions in Autism. *Autism research : official journal of the International Society for Autism Research*. 2016; 9:55–66. [PubMed: 25962831]
- Kleinmans NM, Pauley G, Richards T, Neuhaus E, Martin N, Corrigan NM, et al. Age-related abnormalities in white matter microstructure in autism spectrum disorders. *Brain Research*. 2012; 1479:1–16. [PubMed: 22902768]
- Koldewyn K, Yendiki A, Weigelt S, Gweon H, Julian J, Richardson H, et al. Differences in the right inferior longitudinal fasciculus but no general disruption of white matter tracts in children with autism spectrum disorder. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111:1981–6. [PubMed: 24449864]
- Lee JE, Chung MK, Lazar M, DuBray MB, Kim J, Bigler ED, et al. A study of diffusion tensor imaging by tissue-specific, smoothing-compensated voxel-based analysis. *NeuroImage*. 2009; 44:870–883. [PubMed: 18976713]
- Lerner MD, McPartland JC, Morris JP. Multimodal emotion processing in autism spectrum disorders: An event-related potential study. *Developmental Cognitive Neuroscience*. 2013; 3:11–21. [PubMed: 23245216]
- Ling J, Merideth F, Caprihan A, Pena A, Teshiba T, Mayer AR. Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. *Human Brain Mapping*. 2012; 33:50–62. [PubMed: 21391258]
- Lord, C., Rutter, M., DiLavore, PC., Risi, S. *Manual: Autism Diagnostic Observation Schedule*. Los Angeles, CA: Western Psychological Association; 2002.
- Lunven M, Bartolomeo P. Attention and spatial cognition: Neural and anatomical substrates of visual neglect. *Annals of physical and rehabilitation medicine*. 2016
- Mann TA, Walker P. Autism and a deficit in broadening the spread of visual attention. *Journal of child psychology and psychiatry, and allied disciplines*. 2003; 44:274–284.
- Mapstone M, Weintraub S, Nowinski C, Kaptanoglu G, Gitelman DR, Mesulam MM. Cerebral hemispheric specialization for spatial attention: spatial distribution of search-related eye fixations in the absence of neglect. *Neuropsychologia*. 2003; 41:1396–1409. [PubMed: 12757911]
- Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K, et al. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. *NeuroImage*. 2008; 40:570–582. [PubMed: 18255316]
- Noriuchi M, Kikuchi Y, Yoshiura T, Kira R, Shigeto H, Hara T, et al. Altered white matter fractional anisotropy and social impairment in children with autism spectrum disorder. *Brain Research*. 2010; 1362:141–149. [PubMed: 20858472]
- Oerlemans AM, van der Meer MJ, van Steijn DJ, de Ruiter SW, de Bruijn YGE, de Sonnevile LMJ, et al. Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European child & adolescent psychiatry*. 2014; 23:257–271. [PubMed: 23824472]
- Oishi K, Faria AV, Hsu J, Tippet D, Mori S, Hillis AE. Critical role of the right uncinate fasciculus in emotional empathy. *Annals of Neurology*. 2015; 77:68–74. [PubMed: 25377694]

- Philippi CL, Mehta S, Grabowski T, Adolphs R, Rudrauf D. Damage to association fiber tracts impairs recognition of the facial expression of emotion. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2009; 29:15089–99. [PubMed: 19955360]
- Poustka L, Jennen-Steinmetz C, Henze R, Vomstein K, Haffner J, Sieltjes B. Fronto-temporal disconnectivity and symptom severity in children with autism spectrum disorder. *The World Journal of Biological Psychiatry*. 2012; 13:269–80. [PubMed: 21728905]
- Power JD, Schlaggar BL, Petersen SE. Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage*. 2015; 105:536–551. [PubMed: 25462692]
- Quirk GJ, Beer JS. Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Current Opinion in Neurobiology*. 2006
- Robertson CE, Kravitz DJ, Freyberg J, Baron-Cohen S, Baker CI. Tunnel vision: sharper gradient of spatial attention in autism. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2013; 33:6776–6781. [PubMed: 23595736]
- Rutter, M., Lord, C., LeCouteur, A. *Autism Diagnostic Interview–Revised (ADI–R) manual*. Chicago: Department of Psychiatry, University of Chicago; 1995.
- Sammler D, Grosbras MH, Anwander A, Bestelmeyer PEG, Belin P. Dorsal and Ventral Pathways for Prosody. *Current biology : CB*. 2015; 25:3079–3085. [PubMed: 26549262]
- Schmahmann JD, Pandya DN. *Fiber Pathways of the Brain*. Fiber Pathways of the Brain. 2006
- Schumann CCM, Bloss CSC, Barnes CC, Wideman GM, Carper RA, Akshoomoff N, et al. Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *The Journal of Neuroscience*. 2010; 30:4419–27. [PubMed: 20335478]
- Shen MD, Nordahl CW, Young GS, Wootton-Gorges SL, Lee A, Liston SE, et al. Early brain enlargement and elevated extra-axial fluid in infants who develop autism spectrum disorder. *Brain*. 2013; 136:2825–2835. [PubMed: 23838695]
- Shukla DK, Keehn B, Müller RA. Tract-specific analyses of diffusion tensor imaging show widespread white matter compromise in autism spectrum disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2011; 52:286–295.
- Shukla DK, Keehn B, Lincoln AJ, Müller RA. White matter compromise of callosal and subcortical fiber tracts in children with autism spectrum disorder: a diffusion tensor imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2010; 49:1269–78. 1278–2. [PubMed: 21093776]
- Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006; 31:1487–1505. [PubMed: 16624579]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004; 23
- Solso S, Xu R, Proudfoot J, Hagler DJJ, Campbell K, Venkatraman V, et al. Diffusion Tensor Imaging Provides Evidence of Possible Axonal Overconnectivity in Frontal Lobes in Autism Spectrum Disorder Toddlers. *Biological Psychiatry*. 2014; 79:676–684.
- Sparks BF, Friedman SD, Shaw DW, Aylward EH, Echelard D, Artru AA, et al. Brain structural abnormalities in young children with autism spectrum disorder. *Neurology*. 2002; 59:184–192. [PubMed: 12136055]
- Sundaram SK, Kumar A, Makki MI, Behen ME, Chugani HT, Chugani DC. Diffusion tensor imaging of frontal lobe in autism spectrum disorder. *Cerebral Cortex*. 2008; 18:2659–2665. [PubMed: 18359780]
- Thiebaut de Schotten M, Dell'Acqua F, Valabregue R, Catani M. Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex*. 2012; 48:82–96. [PubMed: 22088488]
- Tournier JD, Mori S, Leemans A. Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*. 2011; 65:1532–1556. [PubMed: 21469191]
- Travers BG, Adluru N, Ennis C, Tromp DPM, Destiche D, Doran S, et al. Diffusion Tensor Imaging in Autism Spectrum Disorder: A Review. *Autism Research*. 2012; 5:289–313. [PubMed: 22786754]

- Unger A, Alm KH, Collins JA, O'Leary JM, Olson IR. Variation in White Matter Connectivity Predicts the Ability to Remember Faces and Discriminate Their Emotions. *Journal of the International Neuropsychological Society : JINS*. 2016; 22:180–90. [PubMed: 26888615]
- Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, et al. Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage*. 2007; 36:630–644. [PubMed: 17481925]
- Wechsler, D. *The Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation; 1999.
- Weinstein M, Ben-Sira L, Levy Y, Zachor DA, Itzhak E Ben, Artzi M, et al. Abnormal white matter integrity in young children with autism. *Human Brain Mapping*. 2011; 32:534–543. [PubMed: 21391246]
- Wolff JJ, Gu H, Gerig G, Elison JT, Styner M, Gouttard S, et al. Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *American Journal of Psychiatry*. 2012; 169:589–600. [PubMed: 22362397]
- Yamada S, Takahashi S, Ukai S, Tsuji T, Iwatani J, Tsuda K, et al. Microstructural abnormalities in anterior callosal fibers and their relationship with cognitive function in major depressive disorder and bipolar disorder: A tract-specific analysis study. *Journal of Affective Disorders*. 2015; 174:542–548. [PubMed: 25556672]
- Yendiki A, Koldewyn K, Kakunoori S, Kanwisher N, Fischl B. Spurious group differences due to head motion in a diffusion MRI study. *NeuroImage*. 2013; 88:79–90. [PubMed: 24269273]

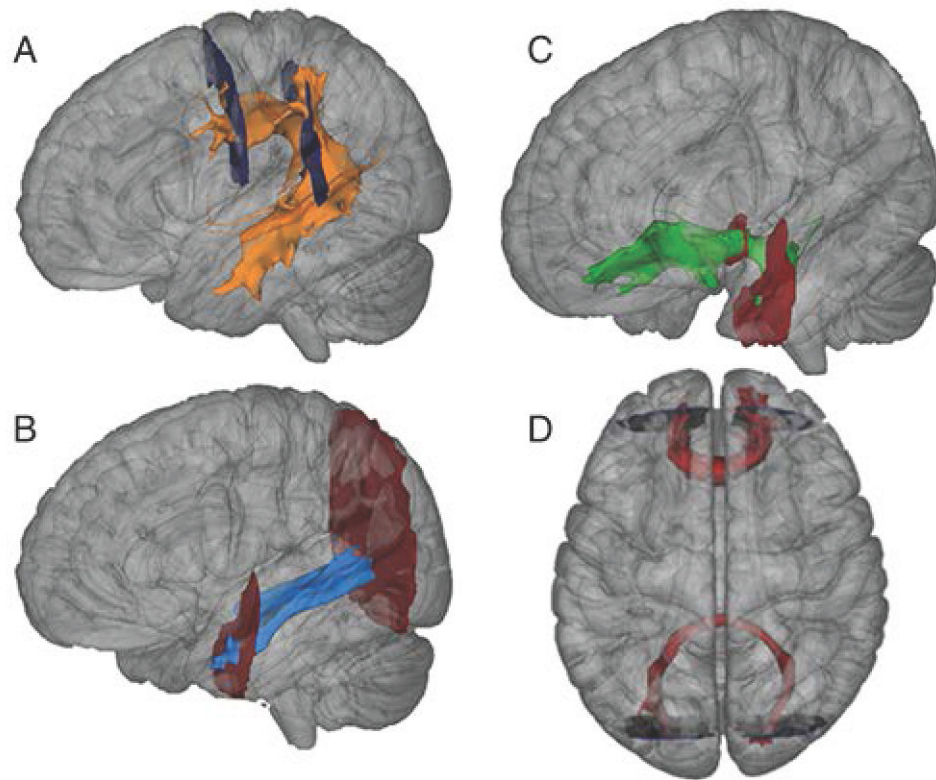


Figure 1.

Tracts of interest for probabilistic tractography analyses. Panels A-C: Tracts (from an exemplary individual TD participant) shown for left hemisphere only. A: Superior Longitudinal Fasciculus. B: Inferior Longitudinal Fasciculus. C: Uncinate Fasciculus. D: Forceps Major (bottom) and Minor (top). Start and stop masks indicated for all tracts.

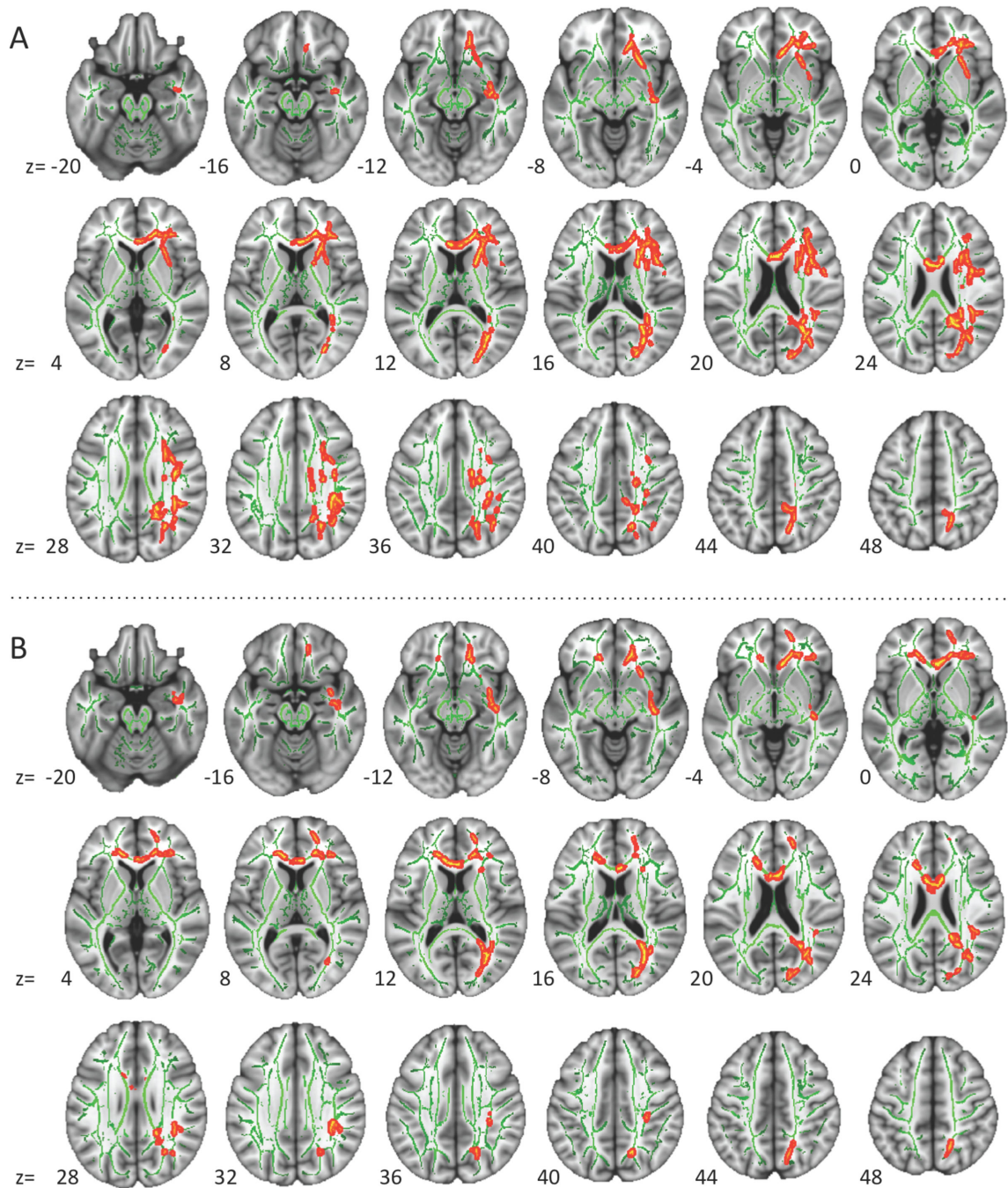


Figure 2.

Significance maps of FA effects from TBSS analyses. A: Before qualitative assessment (FS set; $n_{ASD}=57$, $n_{TD}=50$); B: after quality control (QC set; $n_{ASD}=36$, $n_{TD}=39$). There were no significant effects in the qualitatively matched set. Significant clusters in red-yellow after ‘thickening’ for better visibility; TBSS skeleton in green; right hemisphere displayed to the right.

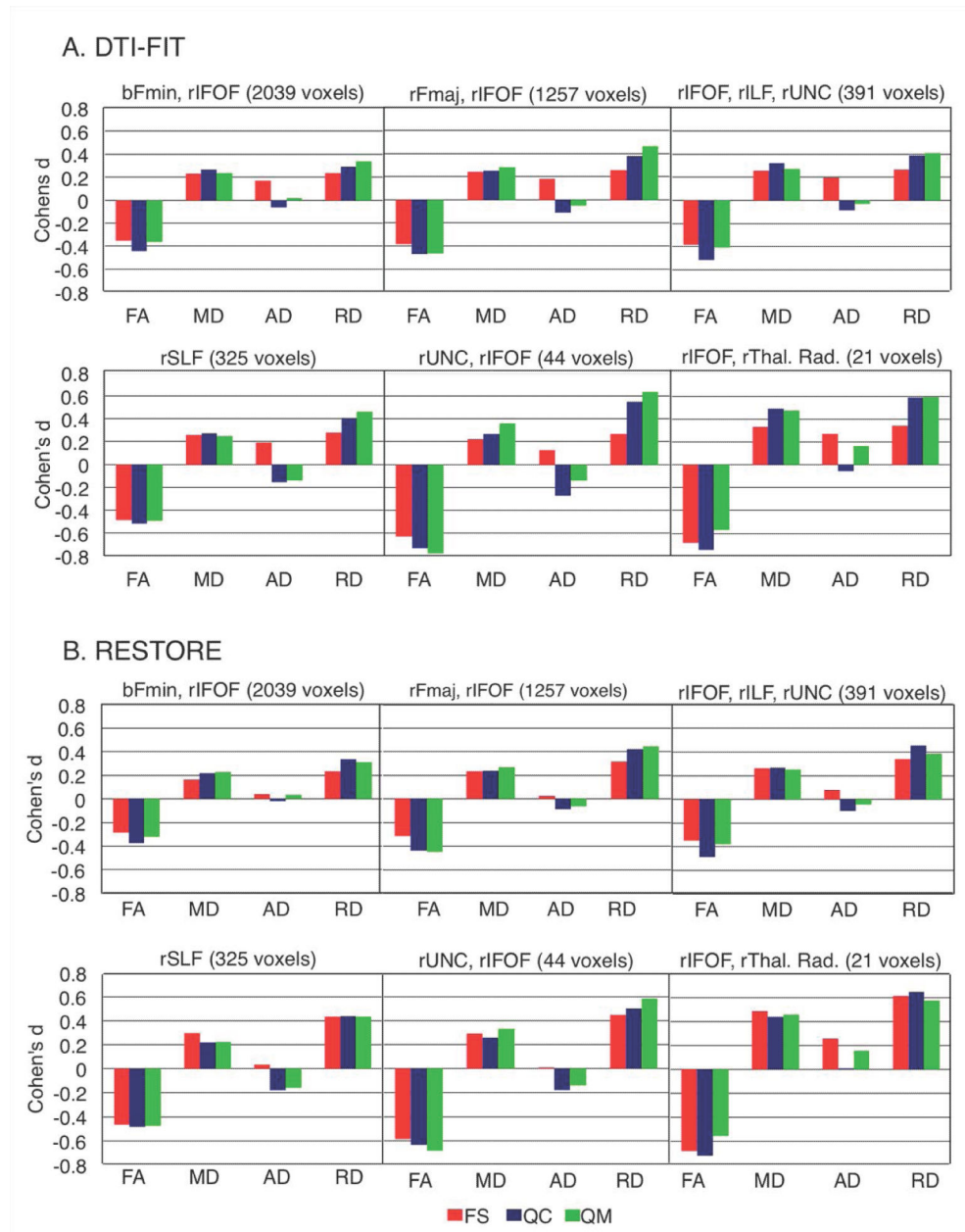


Figure 3.

Effect sizes of group differences in TBSS clusters. Significant clusters from set QC were used as masks to determine average effect sizes (Cohen's d) for FA, MD, AD, and RD in each cluster in each matching stage. A: Effect size results for tensors estimated using *dti-fit*. B: Effect size results for tensors estimated using *restore*. FS = full set; QC = quality controlled; QM = quantitatively matched; rFmin=Right Forceps minor; rIFOF=Right Inferior Fronto-occipital Fasciculus; rSLF=Right Superior Longitudinal Fasciculus; rILF=Right Inferior Longitudinal Fasciculus; b=bilateral; rFmaj=Right Forceps major; UNC=Uncinate Fasciculus; rThal. Rad.=Right thalamic radiation.

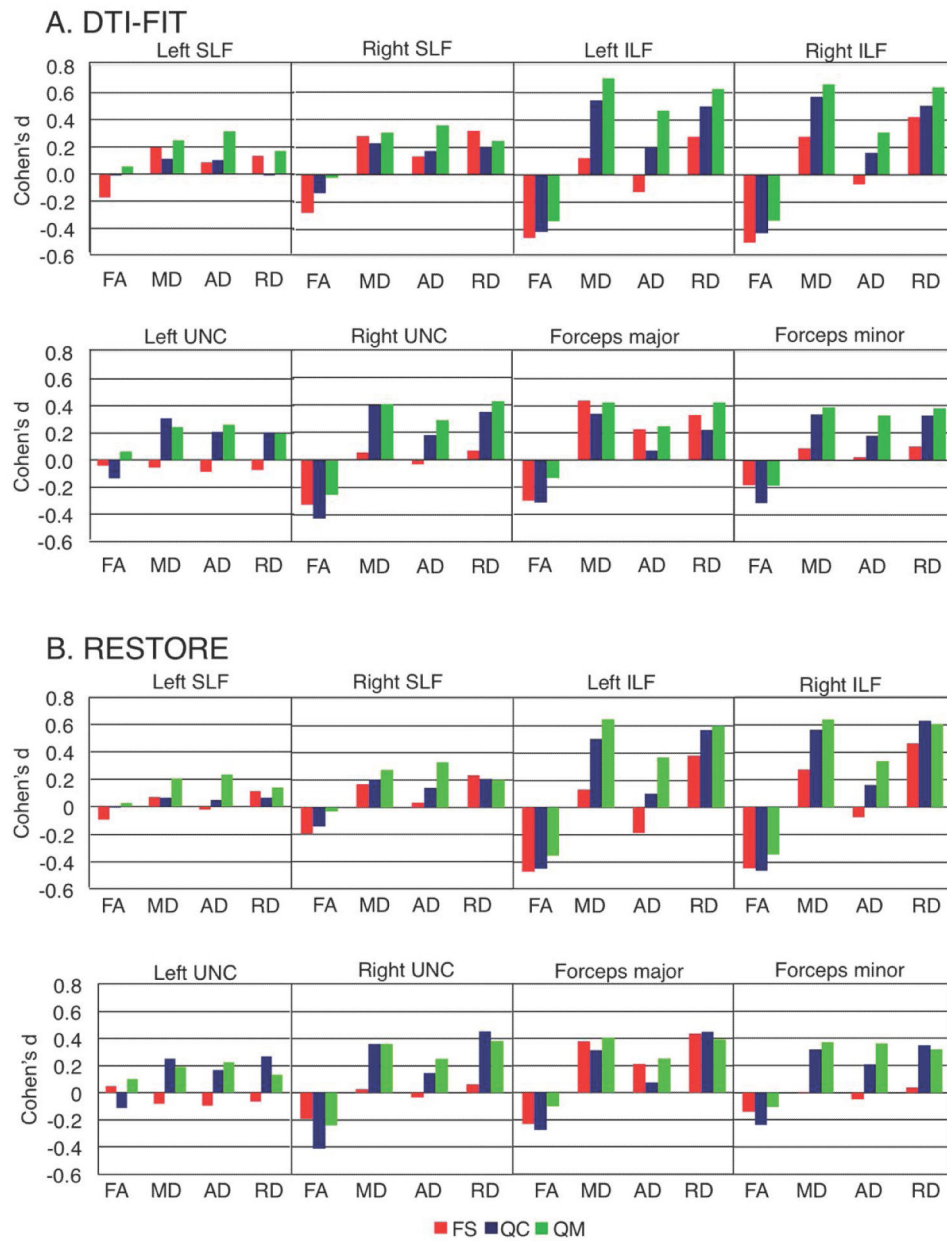


Figure 4. Effect sizes of group differences in tractography. Average effect sizes (Cohen's d) for FA, MD, AD, and RD were calculated within each tract, for each matching stage. A: Effect size results for tensors estimated using *dti-fit*. B.:Effect size results for tensors estimated using *restore*. FS = full set; QC = quality controlled; QM = quantitatively matched; SLF = superior longitudinal fasciculus; ILF = inferior longitudinal fasciculus; UNC = uncinate fasciculus.

Table

Participant demographics and motion measures at each matching stage.

	A. Full Sample (FS set)			B. After Quality Check for Motion (QC set)			C. After Quantitative Motion Matching (QM set)					
	ASD (n=57, 13 females) mean ± SD; [range]	TD (n=50, 12 females) mean ± SD; [range]	Percent Difference	p-value	ASD (n=36, 7 females) mean ± SD; [range]	TD (n=39, 10 females) mean ± SD; [range]	Percent Difference	p-value	ASD (n=30, 6 females) mean ± SD; [range]	TD (n=34, 9 females) mean ± SD; [range]	Percent Difference	p-value
Age (years)	13.28±2.92 [7.64-17.98]	12.75±2.78 [8.01-17.98]	4.02	0.35	13.65±2.88 [8.68-17.80]	12.78±2.63 [8.02-17.59]	6.6	0.17	13.86±3.01 [8.68-17.80]	13.08±2.68 [8.02-17.59]	5.82	0.28
Non-verbal IQ	103.18±18.62 [53-140]	107.40±13.53 [77-137]	4.01	0.19	104.82±19.34 [53-140]	107.28±12.89 [83-137]	2.32	0.52	105.61±17.43 [53-140]	107.76±12.38 [86-137]	2.02	0.57
Verbal IQ	100.41±19.24 [56-147]	107.50±12.60 [73-134]	6.82	0.03	100.12±21.90 [56-147]	106.00±10.97 [73-127]	5.71	0.14	100.68±21.94 [56-147]	107.26±9.89 [87-127]	6.33	0.12
SRS Total	81.00±10.13 [57-100]	42.20±5.46 [35-58]	62.98	<0.001	82.74±9.76 [62-100]	41.97±5.14 [35-52]	65.38	<0.001	82.90±8.81 [62-100]	41.90±5.03 [35-52]	65.7	<0.001
ADOS S+C	12.83±4.15 [7-22]				13.21±4.28 [7-22]				13.74±4.36 [7-22]			
ADOS SBRI	2.30±1.55 [0-6]				2.12±1.43 [0-5]				2.22±1.42 [0-5]			
Average Translation (mm)	1.03±0.49 [0.49-3.03]	0.92±0.34 [0.46-2.38]	11.01	0.46	0.87±0.24 [0.49-1.82]	0.88±0.25 [0.49-1.71]	1.31	0.72	0.84±0.19 [0.49-1.19]	0.89±0.20 [0.49-1.35]	4.93	0.35
Average Rotation (rad; X10 ⁻³)	6.48±4.23 [1.54-21.63]	5.68±3.51 [2.39-20.71]	13.1	0.20	4.86±1.96 [2.31-11.70]	4.64±1.95 [2.39-12.25]	4.73	0.56	4.57±1.55 [2.31-8.56]	4.49±1.59 [2.39-8.63]	1.92	0.78
% Bad Slices	0.13±0.27 [0-1.22]	0.09±0.31 [0-1.99]	39.31	0.03	0.03±0.05 [0.00-0.19]	0.02±0.08 [0.00-0.52]	38.49	0.03	0.02±0.05 [0.00-0.19]	0.02±0.09 [0.00-0.52]	2.14	0.39
Average Dropout Score	1.15±0.19 [1-1.66]	1.07±0.14 [1-1.55]	7.43	0.02	1.15±0.21 [1.00-1.66]	1.03±0.09 [1.00-1.41]	10.47	0.01	1.07±0.13 [1.00-1.38]	1.04±0.10 [1.00-1.41]	3.24	0.34

Group differences in participant demographics and quantitative motion measures at each stage of matching. Age and behavioral measures tested with t-tests, motion measures tested with Mann-Whitney U Test due to non-normal distributions. FS = Full Set. QC = Quality-Checked subsample. QM = Quantitatively Matched subset. SRS = Social Responsiveness Scale. ADOS = Autism Diagnostic Observation Schedule. SBRI = Stereotyped Behaviors and Restricted Interests. S+C = Social and Communication subscores. rad = radians.