Received: 7 July 2021 | Revised: 2 November 2021 | Accepted: 11 November 2021

DOI: 10.1002/hbm.25724

RESEARCH ARTICLE

WILEY

Widespread effects of dMRI data quality on diffusion measures in children

Nabin Koirala[1] | Daniel Kleinman[1] | Meaghan V. Perdue[1,2] | Xing Su[1] | Martina Villa[1,2] | Elena L. Grigorenko[1,3] | Nicole Landi[1,2]

[1]Haskins Laboratories, New Haven, Connecticut, USA

[2]Department of Psychological Sciences, University of Connecticut, Connecticut, USA

[3]Department of Psychology, University of Houston, Houston, Texas, USA

Correspondence
Nabin Koirala, Haskins Laboratories, 300 George Street #900, New Haven, CT 06511, USA.
Email: nabin.koirala@yale.edu

Funding information
Florida Learning Disabilities Research Center, Grant/Award Numbers: NSF DGE-1747453, NSF IGERT DGE-1144399, NIH 2P50HD052120-11

Abstract

Diffusion magnetic resonance imaging (dMRI) datasets are susceptible to several confounding factors related to data quality, which is especially true in studies involving young children. With the recent trend of large-scale multicenter studies, it is more critical to be aware of the varied impacts of data quality on measures of interest. Here, we investigated data quality and its effect on different diffusion measures using a multicenter dataset. dMRI data were obtained from 691 participants (5–17 years of age) from six different centers. Six data quality metrics—contrast to noise ratio, outlier slices, and motion (absolute, relative, translation, and rotational)—and four diffusion measures—fractional anisotropy, mean diffusivity, tract density, and length—were computed for each of 36 major fiber tracts for all participants. The results indicated that four out of six data quality metrics (all except absolute and translation motion) differed significantly between centers. Associations between these data quality metrics and the diffusion measures differed significantly across the tracts and centers. Moreover, these effects remained significant after applying recently proposed harmonization algorithms that purport to remove unwanted between-site variation in diffusion data. These results demonstrate the widespread impact of dMRI data quality on diffusion measures. These tracts and measures have been routinely associated with individual differences as well as group-wide differences between neurotypical populations and individuals with neurological or developmental disorders. Accordingly, for analyses of individual differences or group effects (particularly in multisite dataset), we encourage the inclusion of data quality metrics in dMRI analysis.

KEYWORDS
contrast to noise ratio, diffusion data quality, diffusion measures, motion artifacts, multicenter data

1 | INTRODUCTION

Diffusion magnetic resonance imaging (dMRI) is a noninvasive neuroimaging tool for probing the microstructural architecture of the brain (Basser, Mattiello, & LeBihan, 1994; Merboldt, Hanicke, & Frahm, 1985). In recent years, several dMRI-based models like Diffusion Tensor Imaging (DTI; Basser, Mattiello, & LeBihan, 1994; Pierpaoli, Jezzard, Basser, Barnett, & Di Chiro, 1996), Diffusion Kurtosis Imaging (DKI; Jensen & Helpern, 2010; Lu, Jensen, Ramani, & Helpern, 2006), Diffusion Spectrum Imaging (DSI; Wedeen, Hagmann, Tseng, Reese, & Weisskoff, 2005), q-ball imaging (QBI; Tuch, Reese, Wiegell, & Wedeen, 2003), and high angular resolution diffusion imaging (HARDI; Tuch et al., 2002; Webster & Descoteaux, 2015) have been developed to index white matter alterations (Chanraud, Zahr, Sullivan, & Pfefferbaum, 2010; Koirala et al., 2019), neural density (Koirala, Perdue, Su, Grigorenko, & Landi, 2021; Zhang, Schneider, Wheeler-Kingshott, & Alexander, 2012), and estimates of white matter fiber tracts (Jbabdi & Johansen-Berg, 2011; Koirala et al., 2016). However, as diffusion and the generated indices (e.g., Fractional anisotropy, mean diffusivity) are influenced by several anatomical factors such as fiber arrangements, degree of myelination, cell membranes, microtubules, and axonal integrity, caution must be taken for interpreting the changes associated with these indices (Alba-Ferrara & de Erausquin, 2013; Beaulieu, 2009; Jones & Cercignani, 2010; Jones, Knosche, & Turner, 2013).

Each of these techniques has expanded our understanding of structure–function relationships and various neurological and psychological disorders in both adults and children. Over the last decade, dMRI techniques have been increasingly employed in studies of neurodevelopmental disorders (e.g., attention deficit hyperactivity disorder (ADHD) (Ameis et al., 2016; van Ewijk, Heslenfeld, Zwiers, Buitelaar, & Oosterlaan, 2012; Wu et al., 2017), dyslexia (Vandermosten, Boets, Wouters, & Ghesquiere, 2012; Wang et al., 2017; Yeatman, Dougherty, Ben-Shachar, & Wandell, 2012), autism spectrum disorder (ASD; Andrews et al., 2019; Ismail et al., 2016; Travers et al., 2012), and obsessive–compulsive disorder (OCD; Gruner et al., 2012; Jayarajan et al., 2012; Silk, Chen, Seal, & Vance, 2013, among others). While such studies have been helpful in elucidating differences in microstructural properties in these populations, there has been no systematic evaluation of how variable signal to noise ratio (SNR) and/or the presence of artifact relates to any of the dependent variables of interest, despite the fact that reports consistently note more motion and artifacts among children, especially those with neurodevelopmental disorders (Afacan et al., 2016; Dosenbach et al., 2017; Greene et al., 2018).

In general, compared to other structural imaging techniques, dMRI datasets are often characterized by low signal to noise ratio (SNR; Chilla, Tan, Xu, & Poh, 2015; Polders et al., 2011) and are frequently corrupted by multiple artifacts typically originating from eddy currents, insufficient fat-suppression, B0 inhomogeneity, physiologically related factors (e.g., respiratory motion, cardiac pulsation, participant motion), and Gibbs ringing (Le Bihan, Poupon, Amadon, & Lethimonnier, 2006; Perrone et al., 2015; Pierpaoli, 2012; Tournier, Mori, & Leemans, 2011). Understandably, these artifacts, and particularly motion-related artifacts, are even more significant for the studies involving young children and individuals with neurological developmental disorders (Tamnes, Roalf, Goddings, & Lebel, 2018; Theys, Wouters, & Ghesquiere, 2014). Moreover, it is increasingly common for studies to acquire data from multiple scanners and centers to increase sample size and sample diversity [i.e., Human Connectome Project (Van Essen et al., 2012), Adolescent Brain Cognitive Development study (Casey et al., 2018), etc.]. Pooling data collected from different centers brings new challenges of harmonizing data collection techniques, and data quality from different scanners may be differentially impacted by variations in hardware, scanning sequences, environmental factors, and human factors, as well as other sources of variability (Fortin et al., 2017; Mirzaalian et al., 2016; Ning et al., 2020; Pinto et al., 2020). This is especially relevant when pooling data from existing datasets in which data collection was not optimized for combination with data from other sources. Thus, accounting for cross-scanner differences in image quality when analyzing data collected at multiple centers is a critical step in data analysis.

Over the years, there have been several advances in the attempt to overcome scanning artifacts with different protocols (Bammer, Holdsworth, Veldhuis, & Skare, 2009; Nana, Zhao, & Hu, 2008; Nolte, Finsterbusch, & Frahm, 2000), motion correction techniques (Alhamud, Taylor, Laughton, van der Kouwe, & Meintjes, 2015; Chan et al., 2014; Truong, Chen, & Song, 2011), and better processing algorithms (Andersson et al., 2017; Bastiani et al., 2017; Li et al., 2014). However, there is no consensus on how to compensate for differences in acquisition protocol, artifact handling, data quality control, reconstruction algorithm, visualization approach, or quantitative analysis methodology. Moreover, previous studies have shown that diffusion measures such as the apparent diffusion coefficient (ADC) and fractional anisotropy (FA) vary substantially with different vendors, coil systems, imagers, field strengths, and sequence parameters (Helmer et al., 2016; Sasaki et al., 2008; Schmeel, 2019). The inter-subject, inter-session, inter-site, and between-visits variability were also shown to be significant for these measures with varying degree of significance (Huo et al., 2016; Veenith et al., 2013). In addition, the assumptions made during the acquisition (e.g., perfect field homogeneity, infinitely fast gradient changes, perfectly shaped RF pulses) and processing (e.g., complete correction of susceptibility-induced distortion and eddy current effects, perfect alignment of regions of interests during registration) of DWI data have many pitfalls, and these shortcomings have been shown to effect the computed diffusion measures (Baliyan, Das, Sharma, & Gupta, 2016; Hrabe, Kaur, & Guilfoyle, 2007; Jones & Cercignani, 2010). Even though it is generally accepted that image quality has an impact on the computed measures (Bastin, Armitage, & Marshall, 1998; Soares, Marques, Alves, & Sousa, 2013) and on diffusion measures specifically (Baum et al., 2018; Landman et al., 2007; Ling et al., 2012), research investigating the full range of diffusion measures[1] and the extent of the impact of image quality is still sparse. Here, we investigate (a) several data quality metrics (the ones most relevant to diffusion imaging data: different components of

motion during scanning—absolute, translational, rotational, and relative—as well as contrast to noise ratio and signal drop in different slices, or the number of outlier slices) obtained from children of age range of 5–17 years with varying disability status, and (b) the association of these quality metrics with computed diffusion (fractional anisotropy and mean diffusivity) and tractography (tract density and length) measures obtained from a multicenter dataset with different diffusion modalities and magnetic strength.

2 | METHODS

2.1 | Data acquisition

To assess the relationships between data quality and diffusion measures, we analyzed 708 dMRI data sets of 691 participants who were between 5 and 17 years of age (17 participants in the Italy dataset were scanned twice with two different protocols). Data were obtained from six different centers, using four different models of scanner, five different scanning sequences, two different magnetic strengths (1.5 Tesla [T] and 3 T), and containing two different diffusion models (DTI and DKI). dMRI datasets were obtained from Haskins Laboratories (HL), Boston Children Hospital (BCH), IRCCS Eugenio Medea Italy (ITA), and the Healthy Brain Network Biobank (L. M. Alexander et al., 2017), which includes data from three centers: CitiGroup Cornell Brain Imaging Center (CBIC), Rutgers University Brain Imaging Center (RU) and Staten Island (SI). Sample sizes, demographic information, and data acquisition parameters are presented in Table 1 and Figure 1. All participants included in the study were participating in a multisite study as part of the Florida Learning Disabilities Research Center project VI: *Imaging genetics in SRD: Mega- and meta-analyses*. The data acquisition procedure was approved by the Chesapeake Institutional Review Board (https://www.chesapeakeirb.com/) for the Healthy Brain Network dataset and respective ethics commission for each center. Prior to acquiring the data, written assent obtained from the participant, and written informed consent was obtained from their legal guardians (L. M. Alexander et al., 2017).

2.2 | Reproducibility and data availability

All the raw data (Neuroimaging) used in the study from sites CBIC, RU, and SI are freely available via Healthy Brain Network Biobank (https://childmind.org/center/healthy-brain-network/). The phenotypical data used could be available upon request. The toolbox used in the study (FSL) is an open-access toolbox that can be downloaded from https://fsl.fmrib.ox.ac.uk/fsldownloads_registration. For specific processing protocols used in the study, please refer to Section 2 and contact the corresponding author for any further queries. Scripts used to conduct the analyses reported in this article have been made available on the Open Science Framework (OSF) at https://osf.io/9s27f.

TABLE 1 Details of scanning parameters from all centers

Centers	Number of participants	Mean age (years)	Scanner (field strength)	Diffusion model	Diffusion directions	*b* values (s/mm^2)	Voxel size (mm)
CBIC	228	9.98 ± 2.96	Siemens Prisma (3 T)	DKI	64	0, 1,000, 2,000	2.0 × 2.0 × 2.0
RU	233	9.36 ± 2.86	Siemens TrioTim (3 T)	DKI	64	0, 1,000, 2,000	2.0 × 2.0 × 2.0
SI	117	10.96 ± 3.10	Siemens Avanto (1.5 T)	DKI	64	0, 1,000, 2,000	2.0 × 2.0 × 2.0
HL	37	6.91 ± 0.89	Siemens TrioTim (3 T)	DTI	32	0, 1,000	2.0 × 2.0 × 3.0
BCH	32	5.39 ± 0.27	Siemens TrioTim (3 T)	DTI	30	0, 1,000	2.0 × 2.0 × 2.0
ITA_1	34	13.19 ± 1.94	Philips Achieva (3 T)	DTI	32	0, 300, 1,100	1.7 × 1.7 × 2.0
ITA_2	27	12.90 ± 1.79	Philips Achieva (3 T)	DTI	32	0, 1,100, 2,500	1.7 × 1.7 × 2.0

Note: Here, ITA_1 and ITA_2 is separately shown to highlight the data acquisition difference in scanning parameters (*b* values) used for subanalysis.

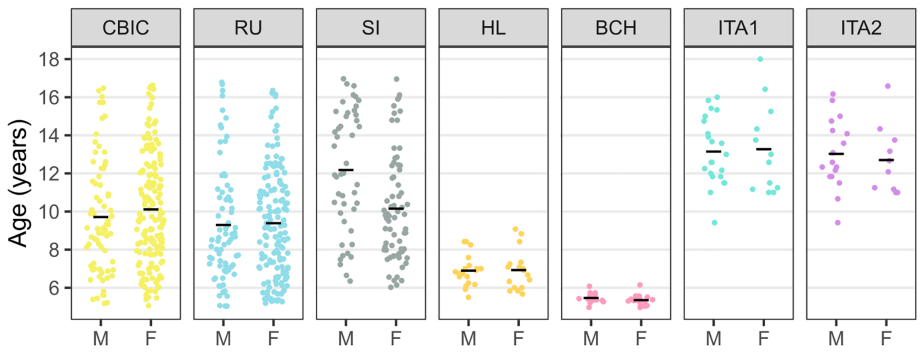

FIGURE 1 Age and sex distribution of all the participants. All participants were selected without limitations on sex or gender, race, or ethnicity, or age other than as scientifically justified. Here M and F indicates male and female and the black dash shows the median line

2.3 | Data analysis

The MRI scans primarily obtained in DICOM format were first converted into Brain Imaging Data Structure (BIDS) format using the *dcm2bids* toolbox available freely at https://github.com/cbedetti/Dcm2Bids (Gorgolewski et al., 2016). Data quality metrics were computed for all participants using an automated quality control framework (named QUAD—QUality Assessment for DMRI; Bastiani et al., 2019) available as an open-source toolbox FSL (ver. 6.0.3). The diffusion data quality metrics (dMRIqc) include average-absolute motion (AAM), average-relative motion (ARM), average-translational motion (ATM), and average-rotational motion (AOM) motion, contrast to noise ratio (CNR), and the number of outlier slices (outlier_dwi). Average absolute motion (with respect to a reference volume) and relative motion (with respect to the previous volume) were calculated as the average voxel displacement across all voxels within a brain mask summarizing both translations and rotations at each voxel (Bastiani et al., 2019). The number of outlier slices indexed how many single slices got distorted during data acquisition because of participant motion that caused signal drop (Andersson & Sotiropoulos, 2016). Here, we computed the total percentage of slices that were distorted within a given volume for each participant.

To obtain the diffusion and tract measures, the images were preprocessed using inbuilt functionality in the open-source toolbox FSL (ver. 6.0.3) described in detail elsewhere (Jenkinson, Bannister, Brady, & Smith, 2002; Jenkinson & Smith, 2001). In brief, data collected with reversed phase-encode blips (except for site BCH) were used for estimating and correcting the susceptibility-induced artifacts using *topup* (Andersson, Skare, & Ashburner, 2003; Smith et al., 2004). Corrections of motion artifacts (eddy currents and head movements) were performed using the *eddy* toolbox (Andersson & Sotiropoulos, 2016), and individual masks were generated for each brain using the *Brain Extraction Toolkit (BET*; Smith, 2002) to isolate the brain from the skull and diffusion tensor modeling for obtaining diffusion measures. Fractional anisotropy (FA) and Mean diffusivity (MD) were calculated using the *FDT* toolbox. In addition, the distribution of crossing fibers was estimated using *BEDPOSTX* (T. Behrens, Berg, Jbabdi, Rushworth, & Woolrich, 2007; T. E. Behrens et al., 2003), and the probability of major (f1) and secondary (f2) fiber directions was calculated (Koirala et al., 2017; Koirala et al., 2019). The obtained crossing fiber modeled diffusion data were further processed using the automatic tractography scheme using *XTRACT* toolbox in FSL (Warrington et al., 2020). Here, tractography masks are defined in standard space; these masks were then warped to each participant's native space using the participant-specific, nonlinear warp fields; and probabilistic tractography was performed in the participant's native space. The resultant tract was then stored in standard space and overlaid on the FSL_HCP1065 FA atlas (Warrington et al., 2020). Finally, tract density, tract length, FA, and MD values were extracted from each of the 36 major tracts as detailed in Table 2.

Using the obtained parameters, statistical analyses were conducted to detect differences between these data quality metrics across centers and brain regions (tracts) and to determine the relationship between these data quality metrics and diffusion measures, as detailed in Section 2.4. In recent years, different harmonization algorithms have been proposed for removing unwanted inter-site variability from multi-site diffusion data (Fortin et al., 2017; Mirzaalian et al., 2016; Ning et al., 2020; Pinto et al., 2020). These algorithms reduce/remove the systematic differences between scanner manufacturers, field strength, and other scanner characteristics that systematically affect the diffusion images and introduce inter-scanner variation. A recent study (Fortin et al., 2017) attempted to address this issue by comparing five statistical harmonization techniques for diffusion data: global scaling, functional normalization (Fortin et al., 2014), removal of artificial voxel effect by linear regression (RAVEL; Fortin et al., 2016), surrogate variable analysis (SVA; Leek & Storey, 2007), and ComBat (Johnson, Li, & Rabinovic, 2007). Of these, the ComBat model, which uses an empirical Bayes (EB) framework to improve the variance of the parameter estimates, was shown to be the most effective in removing unwanted variation induced by site in diffusion data. Accordingly, to evaluate whether applying a harmonization technique in our multi-site dataset would affect the results of our study, we harmonized our data using the ComBat model. To do this, we applied the Matlab-based algorithm (open access availability at https://github.com/Jfortin1/ComBatHarmonization) to harmonize the obtained

TABLE 2 All tracts where the diffusion parameters—tract—density, length, fractional anisotropy (FA), and mean diffusivity (MD) values were extracted

Tract	Abbreviation
Anterior commissure	AC
Arcuate fasciculus (left and right)	Left AF, right AF
Acoustic radiation	AR
Anterior thalamic radiation (left and right)	Left ATR, right ATR
Dorsal cingulum (left and right)	Left CBD, right CBD
Peri-genual cingulum (left and right)	Left CBP, right CBP
Temporal cingulum (left and right)	Left CBT, right CBT
Corticospinal tract (left and right)	Left CST, right CST
Forceps major	FMA
Forceps minor	FMI
Fornix (left and right)	Left FX, right FX
Inferior longitudinal fasciculus (left and right)	Left ILF, right ILF
Inferior fronto-occipital fasciculus (left and right)	Left IFO, right IFO
Middle longitudinal fasciculus (left and right)	Left MdLF, right MdLF
Optic radiation (left and right)	Left OR, right OR
Superior longitudinal fasciculus 1 (left and right)	Left SLF1, right SLF1
Superior longitudinal fasciculus 2 (left and right)	Left SLF2, right SLF2
Superior longitudinal fasciculus 3 (left and right)	Left SLF3, right SLF3
Uncinate fasciculus (left and right)	Left UF, right UF
Vertical occipital fasciculus (left and right)	Left VOF, right VOF

diffusion measures—tract density, tract length, FA, and MD—across sites. A second set of statistical analyses was run using the data harmonized via ComBat.

2.4 | Statistical analysis

Four sets of analyses were conducted in R (v. 3.6.1; R Core Team, 2019) to examine the relationships between six data quality metrics (AAM, ARM, ATM, AOM, CNR, and outlier_dwi) and four diffusion measures (tract density, tract length, FA, and MD) within and across centers. To examine the effects of interest for each set of analyses, post hoc contrasts were applied to the parameters estimated by the statistical models.

2.4.1 | Analysis 1

To determine whether data quality metrics differed by center, six multiple regressions were conducted—one for each data quality metric—to compute effects of center on that metric (the dependent variable). To explore differences in these metrics between centers in more detail, pairwise comparisons were conducted between centers for each metric, with the Tukey method used to control family-wise error rate when the main effect of center was significant.

2.4.2 | Analysis 2

To determine whether diffusion measures differed by center and across brain regions, four linear mixed-effects regression models (Baayen, Davidson, & Bates, 2008) were fit—one for each diffusion measure—to compute effects of center, tract, and their interaction on that diffusion measure (the dependent variable).[2] Contrasts measured the extent to which centers differed in diffusion measures at each of the fiber tracts obtained.

2.4.3 | Analysis 3

To determine the relationship between data quality metrics and diffusion measures as well as any modulation of this relationship by center and brain region, 24 mixed-effects models were fit—one for each combination of data quality metric (6) and diffusion measure (4)—to compute effects of center, tract, data quality metric, and their interactions on that diffusion measure (the dependent variable). Contrasts examined the relationship between each pair of data quality and diffusion measures at each of the fiber tracts obtained.

2.4.4 | Analysis 4

To determine the impact of different scanner models, magnetic strength, diffusion model, and sequence on the relationship between data quality metrics and diffusion measures, we performed contrasts on the statistical models fitted for Analysis 3 (no new models were fit for this analysis). Each contrast compared two centers (or two groups of centers), examining the extent to which the (groups of) centers differed on the relationship between data quality metrics and diffusion measures—both across all tracts, separately at each individual tract. These contrasts examined the effects of several variables:

A. *Scanner models (Siemens Prisma vs. TrioTim)*: By comparing CBIC ($N = 228$, Siemens Prisma) with RU ($N = 233$, Siemens TrioTim).
B. *Magnetic strength (3 vs. 1.5 T)*: By comparing CBIC ($N = 228$, 3 T) and RU ($N = 233$, 3 T) with SI ($N = 117$, 1.5 T).
C. *dMRI models (DKI vs. DTI)*: By comparing CBIC ($N = 228$, DKI) with HL ($N = 37$, DTI) and BCH ($N = 32$, DTI).
D. *dMRI sequence (higher vs. lower b values*[3]*)*: By comparing ITA_1 and ITA_2, where ITA_2 was a partially overlapping subset of participants (overlapping $n = 17$) who were scanned during the same acquisition period but using higher *b* values acquisition sequence (Table 1).

We note that although the centers in each contrast were selected to maximize the relevant difference while holding other variables as constant as possible, all between-center contrasts (A, B, and C) likely captured the joint influences of several variables on diffusion metrics.

All analyses included fixed effects of age and sex, both as main effects and in interactions with all other fixed effects. Including these effects in every model ensured that wherever the dependent variable covaried with age and/or sex, that variance was not attributed to critical predictors (e.g., effects of center or tract or data quality metric). Also, for all analyses, contrast weightings for the effect of center were selected to weight each center in proportion to the number of participants tested at that center (a form of weighted effects coding; cf. [te Grotenhuis et al., 2017]). This means that, for example, when evaluating the main effect of tract, the data from each participant was weighted the same regardless of whether they were tested at center RU ($n = 233$) or center HL ($n = 32$)—which in turn means that, overall, the data from RU was weighted 7.3 times as much as from HL (because 7.3 times as many participants were tested there).

For Analysis 1, which used multiple regressions, significance tests for each factor were conducted via model comparison using the *ANOVA* function. For Analysis 2 and Analysis 3, which used mixed-effects models, significance tests were conducted using the *contestMD* function in the *lmerTest* package (v. 3.1-2; Kuznetsova, Brockhoff, & Christensen, 2017), using the Satterthwaite method to approximate denominator degrees of freedom. (These methods of significance testing were used because they preserved the contrast weights for each center during significance testing.) The mixed-effects models themselves were conducted using the *lme4* package (v. 1.1–21; Bates, Machler, Bolker, & Walker, 2015), employing a two-step model fitting strategy. First, a model was fit with a maximal random effects structure by adding a random intercept for participants, as well

as all within-participant random slopes and their interactions. (Correlations between random slopes were not added because this would have caused the number of random effects parameters to equal the number of observations, making the model unidentifiable. Also, as 674/691 participants (97.5%) were tested at one center each—that is, the entire sample except for the 17 participants tested at the Italy site with two *b* values acquisition sequences—all scans were treated as independent for analytic purposes.) If this model did not converge, all random slopes accounting for less than 1% of the random variance were removed simultaneously (Bates et al., 2015), which always resulted in convergence. Contrasts were computed using the *emmeans* package (v. 1.4; Lenth & Love, 2018). To control Type I error rate, corrections for multiple comparisons were applied. Within each analysis, except as noted below, the Benjamini–Yekutieli method (Benjamini & Yekutieli, 2001) was used to set the false discovery rate (FDR) at 0.05 across all diffusion measures, data quality metrics, and tracts. This correction was applied separately to each effect of interest because each effect represents a different hypothesis. For example, we conducted 24 tests of the hypothesis that there is a main effect of data quality metric on diffusion measures (one test for each combination of six metrics and four measures), so we applied a correction for 24 tests to those *p* values. Separately, we conducted 24 tests of the hypothesis that the effect of data quality metric on diffusion measures interacts with center, so we applied a correction for 24 tests to the *p* values of that two-way interaction. (For each hypothesis test, the number of comparisons corrected for that test is included in the Supporting Information.)

2.5 | Testing the effect of data harmonization

We then ran a second set of analyses on the ComBat-harmonized metrics, including the analyses examining relationships between data quality metrics and diffusion metrics at different centers and tracts. These analyses were identical to our normal analyses—including effects of age and sex on top of the harmonization—except that we also removed the data from center ITA2 to avoid harmonizing data from a subset of participants twice.

3 | RESULTS

3.1 | Analysis 1

We found that four data quality metrics (all except for AAM and ATM) significantly differed between centers, as indicated by a significant main effect of center in four of the six models. Test statistics for each data quality metric are given in Table 3. To explore differences in these metrics between centers in more detail, pairwise comparisons were conducted between centers for each metric, with the Tukey method used to control family-wise error rate when the main effect of center was significant. These pairwise comparisons are shown in Figure 2.

TABLE 3 Details of the results obtained from Analysis 1—To determine whether data quality metrics differed by center

Image quality metric	Between centers
Average relative motion (ARM)	$\mathbf{F_{(6,687)} = 30.2, p < .001}$
Contrast-to-noise ratio (CNR)	$\mathbf{F_{(6,687)} = 25.2, p < .001}$
Outlier slices (outlier_dwi)	$\mathbf{F_{(6,687)} = 13.8, p < .001}$
Average rotational motion (AOM)	$\mathbf{F_{(6,687)} = 4.7, p < .001}$
Average absolute motion (AAM)	$F_{(6,687)} = 0.8, p = 1.00$
Average translation motion (ATM)	$F_{(6,687)} = 0.5, p = 1.00$

Note: The Benjamini–Yekutieli method was used to control FDR across image quality metrics; adjusted *p* values are reported. Significant results are shown in bold.

3.2 | Analysis 2

We found that all analyzed diffusion measures significantly differed between centers and between tracts, and that these effects interacted. Test statistics for each diffusion measure are presented in Table 4. Descriptively, centers differed most strongly on MD, followed by tract density, tract length, and FA (Figure 3), whereas tracts differed most strongly on tract length and FA, followed by MD and tract density. All 36 tracts significantly differed across centers on FA, MD, and tract length; and 11/36 tracts significantly differed across centers on tract density (Figure 4).

3.3 | Analysis 3

Associations between diffusion measures and data quality metrics varied across tracts to different degrees at different centers. This was indicated by significant three-way interactions between center, tract, and data quality metrics for 13/24 (54%) combinations of diffusion measures and data quality metrics (Figure 5, bottom row). Relationships with MD were the most influenced, showing significant three-way interactions with center and tract on 5/6 data quality metrics, followed by tract density (4/6 data quality metrics). The least impacted relationships were with FA and tract length, for which center and tract jointly modulated relationships with 2/6 data quality metrics each (both average CNR and ARM). Relatedly, we found that relationships with ARM were significantly modulated by center and tract for 4/4 diffusion measures, and relationships with CNR were significantly modulated by center and tract for 3/4 diffusion measures (all except tract density).

Although the presence of three-way interactions qualifies the interpretation of two-way interactions, we note for completeness that for MD and tract length, relationships with several data quality metrics significantly differed across both centers and tracts. For the remaining diffusion measures, however, significant relationships with one or more data quality metrics were observed either across centers (FA) or across tracts (tract density; see Figure 5, middle rows).

A small number of significant relationships (11/864, after correcting for multiple comparisons) were observed between diffusion

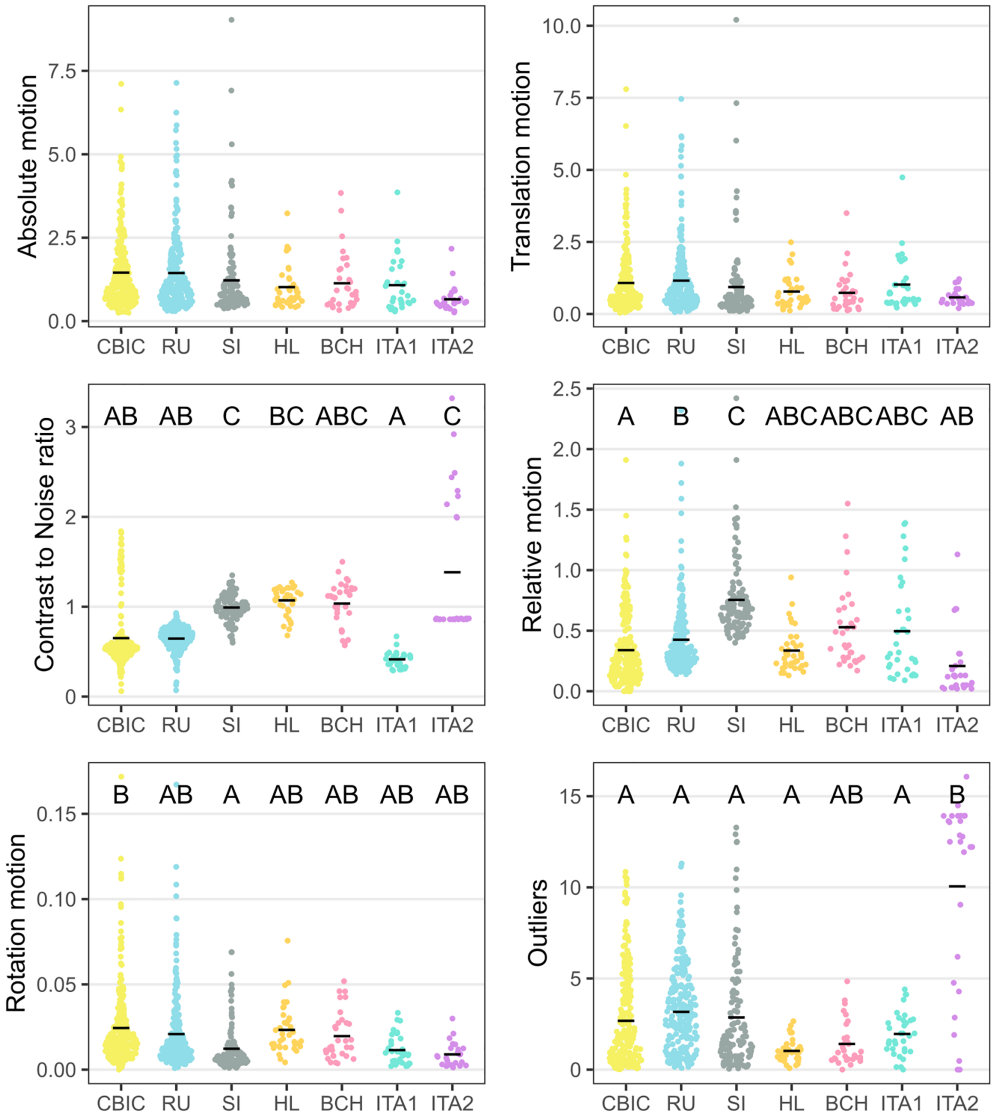

FIGURE 2 Data quality metrics for all centers. For metrics that significantly differed across centers, letters above each center show the results of Tukey-corrected pairwise comparisons: Centers that do not share a letter significantly differed on the metric after regressing out age and sex. For example, center CBIC significantly differed from centers SI and ITA2, but not any other center, on contrast-to-noise ratio

Diffusion measure	Between centers	Between tracts	Center × tract
FA	$F_{(6,687)} = 17.7$, $p < .001$	$F_{(35,1,352)} = 136.4$, $p < .001$	$F_{(210,1,352)} = 35.7$, $p < .001$
MD	$F_{(6,687)} = 595.0$, $p < .001$	$F_{(35,1,604)} = 71.6$, $p < .001$	$F_{(210,1,604)} = 48.0$, $p < .001$
Tract density	$F_{(6,687)} = 188.3$, $p < .001$	$F_{(35,2,746)} = 53.3$, $p < .001$	$F_{(210,2,746)} = 99.3$, $p < .001$
Tract length	$F_{(6,687)} = 64.0$, $p < .001$	$F_{(35,1,159)} = 170.6$, $p < .001$	$F_{(210,1,159)} = 37.9$, $p < .001$

TABLE 4 Details of the results obtained from Analysis 2—To determine whether diffusion measures differed by center and across brain regions

Note: The Benjamini–Yekutieli method was used to control FDR across diffusion measures; adjusted *p* values are reported. Significant results are shown in bold.

measures and image quality metrics *within* individual tracts. Positive relationships were observed between CNR and MD in right optic radiation, left superior longitudinal fasciculus 1, and bilateral uncinate fasciculus; and between CNR and tract length in left corticospinal tract, left inferior fronto-occipital fasciculus, right superior longitudinal fasciculus 1, left superior longitudinal fasciculus 2, and bilateral uncinate fasciculus; in addition, a negative relationship was observed between the number of outlier slices and MD in the forceps major. The results obtained from these analyses provide further evidence that the relationships between data quality metrics and diffusion measures are irregularly modulated across different tracts (even if they are only significant at a few individual tracts) and vary significantly across centers.

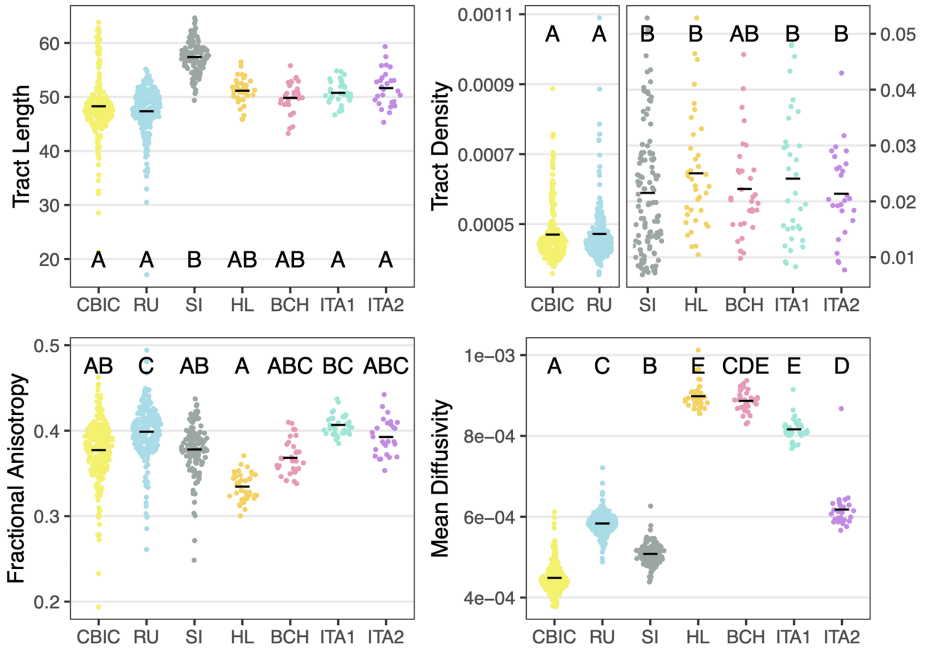

FIGURE 3 Diffusion measures for all centers. For measures that significantly differed across centers, letters above each center show the results of Tukey-corrected pairwise comparisons: Centers that do not share a letter significantly differed on the measure after regressing out age and sex. For example, center CBIC significantly differed from center SI, but not any other center, on tract length. As analyses controlled for participant age and sex, centers that were outliers in their distribution of participant ages (e.g., center BCH) may show fewer significant differences with other centers here than their means would suggest

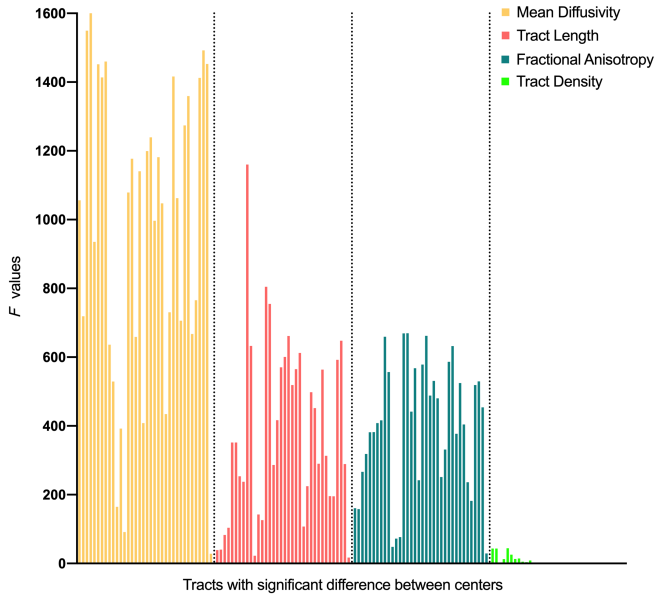

FIGURE 4 All tracts that differ significantly across centers on Mean diffusivity (MD) (36/36), fractional anisotropy (36/36), tract length (36/36), and tract density (11/36). Each line indicates a tract with the height of it showing strength. Please refer to Supporting Information for the details of all tracts shown here with the statistical values

3.4 | Analysis 4

3.4.1 | Scanner models (Siemens Prisma vs. TrioTim)

When comparing centers differing in the scanner manufacturer (CBIC vs. RU), we found several significant differences in the relationships between diffusion measures and data quality metrics, both across tracts and at individual tracts. Modulation across tracts was observed for the relationship between MD and image quality metrics (ARM, ATT, and CNR), as well as the relationships between FA and CNR, and between tract length and ARM (Table 5). Moreover, the relationships between the metrics varied between tracts, with the largest number of significant effects (22) associated with MD, then tract length (6) and FA (5) [and none with tract density]. All tracts at which a significant difference was observed between centers in the relationship between diffusion measures and data quality metrics are shown in Figure 6.

3.4.2 | Magnetic strength (3 vs. 1.5 T)

When comparing centers differing in magnet strength (CBIC and RU vs. SI), no significant differences were observed when averaging across tracts. Comparing centers at each tract, all three significant differences were in the relationship between ARM and various diffusion metrics in the temporal cingulum, with stronger relationships observed for the 1.5 T scanner. The details of all tracts showing significant differences are shown in Figure 6. (Note that the centers in this contrast also varied on scanner model, as shown in Table 1.)

3.4.3 | dMRI models (DKI vs. DTI)

When comparing centers that differed in the diffusion models (DKI and DTI) used to acquire data (CBIC vs. HL and BCH), no significant differences were observed when averaging across tracts. Comparing centers at each tract, one significant difference was observed in the relationship between outlier_dwi and MD at FMA, with a stronger

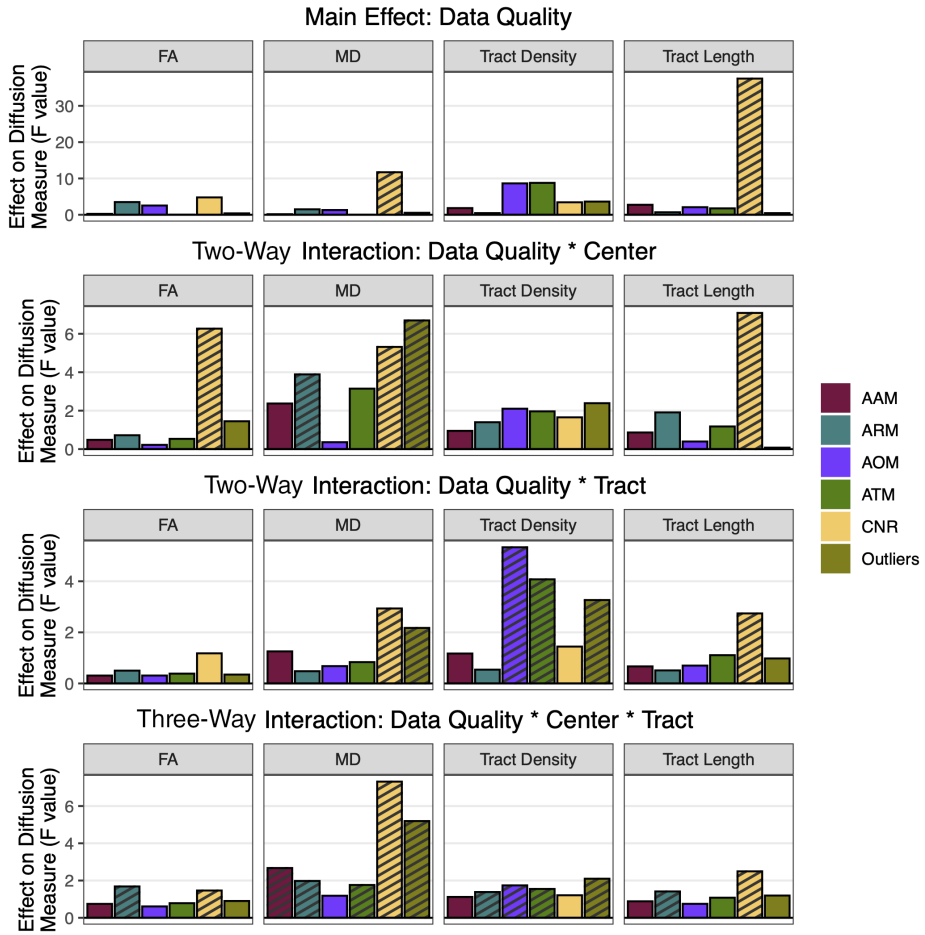

FIGURE 5 Figure visualizing Analysis 3 effects of data quality metrics on diffusion measures. Each row corresponds to a different type of effect (main effect of a data quality metric in the top row, two-way interactions with center or tract in the middle rows, three-way interaction with both center and tract in the bottom row). Each column corresponds to a different diffusion measure. Within each panel, each colored bar corresponds to a different data quality metric. The *Y*-axis represents the *F*-value of the given effect with the given data quality metric on the given diffusion measure. Effects that were statistically significant after correction for multiple comparisons are shaded with diagonal slashes. (Bar heights are directly comparable within a row but are not comparable between rows due to differences in degrees of freedom for different effects.) As an example of how to interpret the figure, the prevalence of many shaded bars in the MD column means that multiple significant relationships were observed between data quality metrics and MD, whereas the prevalence of many yellow shaded bars throughout the figure means that CNR was significantly related to multiple diffusion measures in different ways. Please refer to Supporting Information for the details of all tracts shown here with the statistical values. AAM, average absolute motion; AOM, average rotational motion; ARM, average relative motion; ATM, average translational motion; CNR, contrast to noise ratio; FA, fractional anisotropy; MD, mean diffusivity

relationship observed for the DTI model. (Note that the centers in this contrast also varied on scanner model, the number of diffusion directions, *b* values, and voxel size, as shown in Table 1.)

3.4.4 | dMRI sequence (higher vs. lower *b* values)

For this dataset acquired during the same session with two different *b* values (ITA1 vs. ITA2), we did not observe any significant differences when averaging across tracts, but significant differences did emerge in the relationships between data quality metrics and two diffusion measures—MD and tract density—at a small number of tracts (Figure 6). These effects were not always more prominent for the higher *b* values sequence (probably as one might assume) but were rather dependent on the data quality metric and diffusion measures.

3.4.5 | The effect of harmonization

The analyses with ComBat-harmonized data revealed that the harmonization did not eliminate or even reduce the number of significant relationships observed between the six data quality metrics and the four diffusion measures. In Analysis 3, those relationships

a. were significant (main effect) for 4/24 combinations of metrics (vs. 2/24 for non-harmonized data);

TABLE 5 Details of significant ($p < .05$) results from Analysis 4—To determine the impact of different scanner models, magnetic strength, diffusion model, and sequence to the relationship of data quality metrics and diffusion measures

Scanner models		
Diffusion measure	**Data quality metric**	**B ± SE**
FA	CNR	−0.02 ± 0.004
MD	ARM	1E−5 ± 3E−6
MD	ATM	1E−5 ± 3E−6
MD	CNR	1E−5 ± 5E−6
Tract length	ARM	1.64 ± 0.513

Note: The Benjamini–Yekutieli method was used to control FDR across diffusion measures. Here, B indicates the Beta coefficient representing the mean effect of the predictor (data quality metrics) on the dependent variable (DV; diffusion measures) for one center minus the effect of the predictor on the DV for another center. The overall contrast for this analysis was only significant for scanner models. See full results provided at https://osf.io/9s27f.

Abbreviations: ARM, average relative motion; ATM, average translation motion; CNR, contrast to noise ratio; FA, fractional anisotropy; MD, mean diffusivity; SE, standard error.

b. significantly varied by center (two-way interaction) for 0/24 combinations of metrics (vs. 5/24 for non-harmonized data);
c. significantly varied by tract (two-way interaction) for 5/24 combinations of metrics (vs. 6/24 for non-harmonized data); and
d. significantly varied jointly by center and tract (three-way interaction) for 17/24 combinations of metrics (vs. 13/24 for non-harmonized data)

Regarding tract-by-tract analyses, of all combinations of 36 tracts, six data quality metrics, and four diffusion metrics, the relationship between metrics was significant at an individual tract for 12/864 combinations (vs. 11/864 for non-harmonized data).

In general, ComBat harmonization did appear to reduce between-center differences in diffusion metrics (the main effect of center and the interaction between center and tract were each significant for 0/4 metrics in Analysis 2, vs. 4/4 for non-harmonized data). However, perhaps surprisingly, the number of significant interactions with age and sex in all analyses was roughly comparable for harmonized and non-harmonized data.

All the statistical results from the Main as well as from ComBat analysis have been provided as Supporting Information in the OSF platform (https://osf.io/9s27f).

4 | DISCUSSION

In this study, involving a large multisite dataset of typically developing children and children with neurodevelopmental disorders, we demonstrated that the diffusion imaging data quality, assessed via multiple metrics, differed significantly between data collection sites. Furthermore, the associations between these data quality indices and the obtained diffusion measures differed non-uniformly across sites and brain areas. Analyzing different components of data acquisition, we found significant modulation of the relationship between data quality metric and diffusion measures were significantly influenced by scanner models, magnetic strength, diffusion models, and data collection sequences.

Over the last two decades, dMRI has been established as an effective technique for understanding various neurodevelopmental conditions and disorders and psychological states and processes. Several indices have been proposed encompassing a wide range of microstructural properties including anisotropy, diffusivity, fiber orientation, and neurite features. These indices are known to be influenced by several white matter properties such as fiber arrangements, degree of myelination, and axonal integrity (Alba-Ferrara & de Erausquin, 2013; Jones et al., 2013). Moreover, the contributions of these white matter properties to computed diffusion measures are non-linear and not well known. To give an example, in a recent study using animal models treated with different drugs to trigger demyelination and axonal damage, the FA values obtained could not distinguish between a demyelination group and those with both demyelination and axonal damage (Boretius et al., 2012). Similar concerns have been raised for crossing fibers (around 90% of white matter voxels contain crossing fibers) and q-space (diffusion space), which have been shown to have a remarkable impact on fiber tract estimation and anisotropy analysis (Jeurissen, Leemans, Tournier, Jones, & Sijbers, 2013; Wilkins, Lee, Gajawelli, Law, & Lepore, 2015). The susceptibility of dMRI data to various artifacts (e.g., eddy currents, echo-planar distortions, rotation errors, partial volume effects, and scanner artifacts) has been studied to better understand how these artifacts impact the estimation of eigenvalues and eigenvectors, which in turn influence the accuracy of computed anisotropic measures and fiber tracking schemes (Anderson, 2001; Skare, Li, Nordell, & Ingvar, 2000). Furthermore, one of the largest confounds for studies in children and adolescents—head motion—has been shown to have a significant impact on results obtained in group comparison studies (Yendiki, Koldewyn, Kakunoori, Kanwisher, & Fischl, 2014) or in studies investigating the relationship to diffusion measures (Baum et al., 2018). Even though several studies have highlighted these aspects and proposed new methods (Li et al., 2014; Oguz et al., 2014), the consideration of data quality for the analysis and a transparent, standardized estimate of quality assurance is still rare. Most of the studies suggest either removing the artifacts during preprocessing steps or manually performing quality control as a good practice for controlling for this issue. In this study, we provide evidence that even after these standard preprocessing steps for artifacts correction, the relationships between these data quality measures and computed diffusion metrics are still widespread across the brain and vary non-uniformly across tracts and data collection centers. This result has potentially important implications for the inferences made by the clinical and neurobiological studies that use these metrics.

The results from Analysis 1 and Analysis 2 illustrated that data quality metrics (except for absolute and translational motion) and computed diffusion measures varied across centers. This provides one possible scientific explanation for differences in findings and

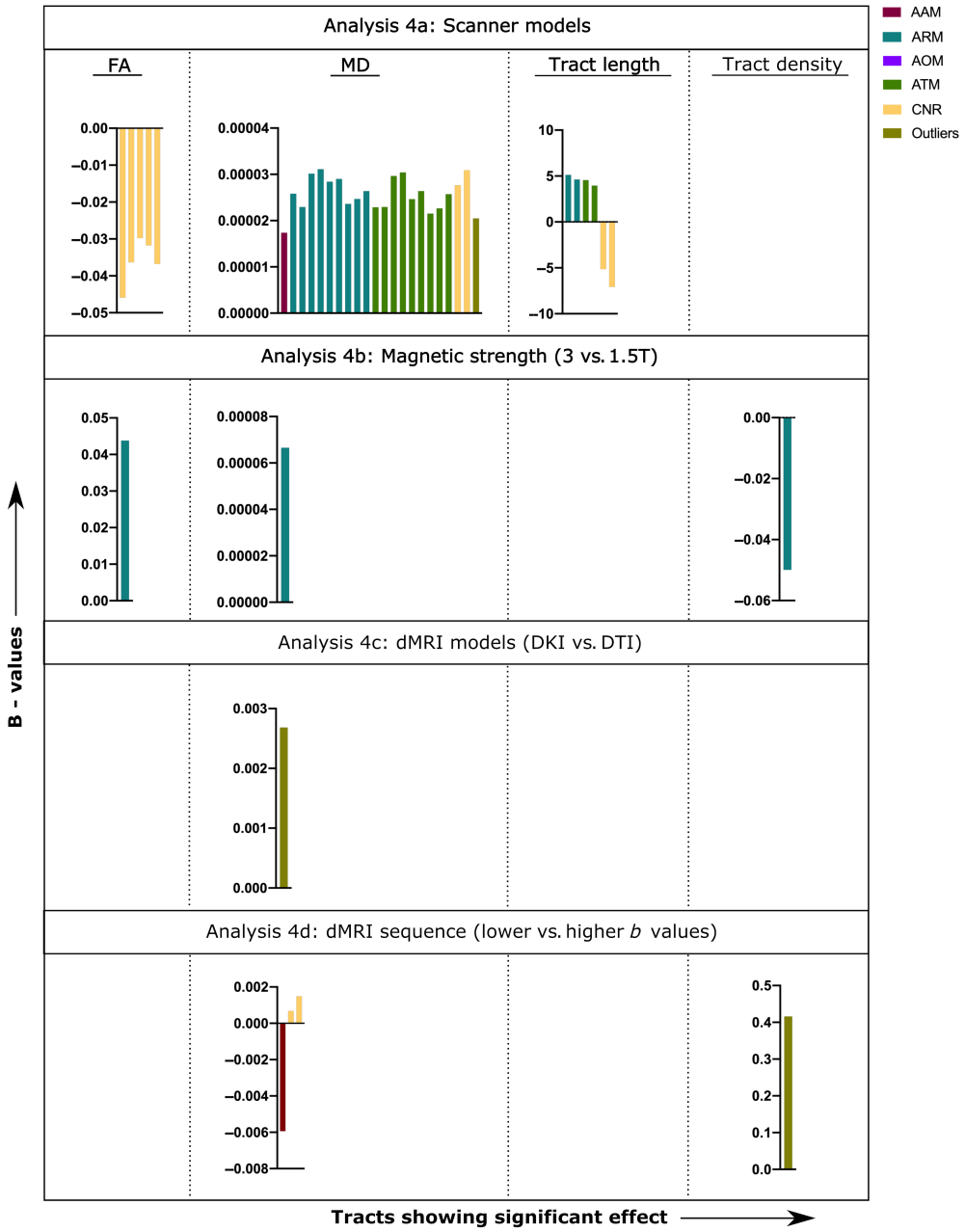

FIGURE 6 All tracts that showed statistically significant differences between the centers in the effects of each data quality metric in each tract on each diffusion measures. Please refer to Supporting Information for the details of all tracts shown here with the statistical values. Here, B indicates the Beta coefficient representing the mean effect of the predictor (data quality metrics) on the dependent variable (DV) (diffusion measures) for one (group of) center(s) minus the effect of the predictor on the DV for another (group of) center(s). AAM, average absolute motion; AOM, average rotational motion; ARM, average relative motion; ATM, average translational motion; CNR, Contrast to noise ratio; FA, fractional anisotropy; MD, mean diffusivity

replication problems for studies conducted across different centers. Moreover, we observed that the impact of these data quality metrics on diffusion measures varies across centers and across different tracts in the brain (Analysis 3). The tracts which showed the strongest variation across centers for different diffusion measures are commonly studied in clinical populations and in relation to various neuropsychological parameters. In addition, we observed that MD, tract density, tract length, and FA, which are some of the widely used parameters in neurodevelopmental studies, were the parameters highly influenced by the data quality metrics. However, to the best of our knowledge, no prior neurodevelopmental study has accounted for these data quality metrics in the statistical models. A failure to do so might lead to unaccounted impacts on suboptimal tensor estimation, erroneous partial volume effect computation, or estimation of boundary threshold. Accordingly, we recommend computing these quality metrics and including them in analyses when reporting delicate microstructural changes.

In Analysis 4, the four subanalyses completed to quantify these differences across different scanner models, magnetic strength, diffusion models, and q-space sampling or *b* values, revealed significant modulation of the relation between quality metrics and diffusion measures by these factors. For these comparisons, MD was the measure that showed the largest number of significant relationships with data quality measures, and ARM was the data quality metric that showed the largest number of significant relationships with diffusion measures. More importantly, the impact of data quality varied (e.g., direction of association,: positive or negative and strength of the association) across different tracts and different diffusion measures,

making it impossible to process out one or more effects. For example, the effect of CNR in FA values is higher in most tracts for *Siemens TrimTrio*; however, the effect of the same in MD is higher for *Siemens Prisma*. In addition, the strength of the relationship between the FA, MD, tract length and tract density is not uniform for each of the data quality metrics and for each tract (refer to Figure 6 for visualization of these examples). This variation further highlights the importance of using the quality metrics in statistical analysis while computing these diffusion measures to capture this unseen influence. Moreover, for multisite data, this might be a promising way to gauge differences in the datasets and evaluate them for harmonization.

After harmonizing our data across centers using the ComBat algorithm, we continued to observe significant relationships between data quality metrics and diffusion metrics, both directly and varying across centers and tracts. Of these effects, the total number of significant effects was the same with vs. without harmonization (26/96 significant). This further supports the necessity of considering data quality metrics in data analysis even after harmonization.

In recent years, there have been several studies which have highlighted the impact of quality assurance of dMRI data on computed measures (Maximov, Alnaes, & Westlye, 2019; Roalf et al., 2016) and have suggested various optimization methods (Liu et al., 2019) and alternative diffusion measures less sensitive to the data quality (Ozcan, 2010). With this study, we presented quantitative sets of data quality metrics (that can be straightforwardly obtained using standard toolboxes) and comprehensive details of their impact on the computed diffusion measures. We suggest that moving forward, these quality metrics should be considered in data analysis, particularly for multisite studies and those involving populations prone to movement. The dataset used for this analysis was aggregated post-hoc; thus, data collected at each center consisted of a different set of participants. Moreover, participant diagnoses for different neurodevelopmental and neuropsychological disorders were not modeled during analysis as this was outside the scope of the study. Although it is possible that including diagnosis could change which specific relationships with data quality metrics were statistically significant, we think that the overall pattern of results—that is, the existence of significant relationships between these metrics and diffusion measures, often modulated by center and/or tract—would remain the same. Despite the aggregated nature of the dataset, we showed strong effects of center and influence of different components involved in data acquisition including scanner model, magnetic strength, and diffusion models (though some of these comparisons may reflect other between-center variables as well). We conclude that this assessment of image quality in such a dataset is important as it most closely represents the types of large-scale multisite studies that researchers currently draw upon for conducting large-scale neuroimaging research. However, ongoing and future multisite studies can and should further optimize harmonization across sites through the use of concurrent acquisition models and sequence where possible, evaluate the difference with the use of phantoms beforehand, and use some of the harmonization techniques (Fortin et al., 2017; Mirzaalian et al., 2016; Ning et al., 2020) for closely related sequences.

5 | CONCLUSION

We provided evidence that several dMRI data quality metrics differ between imaging data collection sites. More importantly, the relationship between these data quality metrics and diffusion measures are widespread across brain regions, though this too varied across center. The obtained brain areas and diffusion measures have been widely employed to discriminate neurotypical populations from individuals with neurological or developmental disorders. Hence, these data quality metrics should be considered for inclusion in dMRI analysis, which could be quantified using available toolboxes. Taking these metrics into account would help make data more comparable across testing sites, would provide more reliable findings, and might be a key in producing replicable studies for group comparison, identifying correlates of individual differences, and carrying out large-scale multisite studies to identify neural markers for different neurodevelopmental conditions and cognitive functions.

ACKNOWLEDGMENTS

The work was supported by Florida Learning Disabilities Research Center, Grant/Award Number NIH 2P50HD052120-11 to Florida State University (PR: Richard Wagner) and Grant/Award Number NSF IGERT DGE-1144399 to the University of Connecticut. Co-author Meaghan V. Perdue was supported by NSF DGE-1747453. We would like to express our sincere gratitude to the collaborators who provided us the data for the analysis, Michael P. Milham, Child Mind Institute (Healthy Brain Network); Filippo Arrigoni, Denis Peruzzo, and Sara Mascheretti, Scientific Institute, IRCCS Eugenio Medea, Italy; Nadine Gaab, Boston children's hospital, and Kenneth R. Pugh, Haskins Laboratories.

CONFLICT OF INTEREST

All authors have no conflict of interest to declare.

ENDNOTES

[1] All diffusion measures were computed based on the principle of diffusion tensors, which describes the covariance of diffusion displacements in three dimensions normalized by the diffusion time. The diffusion coefficient (the primary index in this model) is a measure of the magnitude of diffusion (of water molecules) within tissue, expressed in mm^2/s, and is proportional to the mean-squared displacement divided by the number of dimensions and the diffusion time. For mathematical details, please refer to A. L. Alexander, Lee, Lazar, and Field (2007), Basser et al. (1994), Einstein, Fürth, and Cowper (1926) and Pierpaoli et al. (1996).

[2] As diffusion measures are always computed for individual tracts, main effects of center represent an overall shift in diffusion measure values "across tracts"; that is, with each tract represented equally, though significant interactions between center and tract would indicate that the extent of this shift varied by tract.

[3] b values measure the degree of diffusion weighting applied (in s/mm^2), thereby indicating the amplitude (G), time of applied gradients (δ), and duration between the paired gradients (Δ), formulated as: $b = \gamma^2 G^2 \delta^2 (\Delta - \delta/3)$ (Brown, Cheng, Haacke, Thompson, & Venkatesan, 2014).

DATA AVAILABILITY STATEMENT

All the raw data (Neuroimaging) used in the study from site CBIC, RU and SI are freely available via Healthy Brain Network Biobank

(https://childmind.org/center/healthy-brain-network/). The phenotypical data used could be available upon request. The toolbox used in the study (FSL) is an open access toolbox which can be downloaded from https://fsl.fmrib.ox.ac.uk/fsldownloads_registration. For specific processing protocols used in the study, please refer to the Methods section and contact the corresponding author for any further queries. Scripts used to conduct the analyses reported in this paper have been made available on the Open Science Framework (OSF) at https://osf.io/9s27f.

ORCID

Nabin Koirala https://orcid.org/0000-0002-8261-8271

REFERENCES

Afacan, O., Erem, B., Roby, D. P., Roth, N., Roth, A., Prabhu, S. P., & Warfield, S. K. (2016). Evaluation of motion and its effect on brain magnetic resonance image quality in children. *Pediatric Radiology*, *46*(12), 1728–1735. https://doi.org/10.1007/s00247-016-3677-9

Alba-Ferrara, L. M., & de Erausquin, G. A. (2013). What does anisotropy measure? Insights from increased and decreased anisotropy in selective fiber tracts in schizophrenia. *Frontiers in Integrative Neuroscience*, *7*, 9–9. https://doi.org/10.3389/fnint.2013.00009

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

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., … Milham, M. P. (2017). An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data*, *4*, 170181. https://doi.org/10.1038/sdata.2017.181

Alhamud, A., Taylor, P. A., Laughton, B., van der Kouwe, A. J. W., & Meintjes, E. M. (2015). Motion artifact reduction in pediatric diffusion tensor imaging using fast prospective correction. *Journal of Magnetic Resonance Imaging*, *41*(5), 1353–1364. https://doi.org/10.1002/jmri.24678

Ameis, S. H., Lerch, J. P., Taylor, M. J., Lee, W., Viviano, J. D., Pipitone, J., … Anagnostou, E. (2016). A diffusion tensor imaging study in children with ADHD, autism Spectrum disorder, OCD, and matched controls: Distinct and non-distinct white matter disruption and dimensional brain-behavior relationships. *The American Journal of Psychiatry*, *173*(12), 1213–1222. https://doi.org/10.1176/appi.ajp.2016.15111435

Anderson, A. W. (2001). Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magnetic Resonance in Medicine*, *46*(6), 1174–1188. https://doi.org/10.1002/mrm.1315

Andersson, J. L. R., Graham, M. S., Drobnjak, I., Zhang, H., Filippini, N., & Bastiani, M. (2017). Towards a comprehensive framework for movement and distortion correction of diffusion MR images: Within volume movement. *NeuroImage*, *152*, 450–466. https://doi.org/10.1016/j.neuroimage.2017.02.085

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20*(2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

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

Andrews, D. S., Lee, J. K., Solomon, M., Rogers, S. J., Amaral, D. G., & Nordahl, C. W. (2019). A diffusion-weighted imaging tract-based spatial statistics study of autism spectrum disorder in preschool-aged children. *Journal of Neurodevelopmental Disorders*, *11*(1), 32. https://doi.org/10.1186/s11689-019-9291-z

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baliyan, V., Das, C. J., Sharma, R., & Gupta, A. K. (2016). Diffusion weighted imaging: Technique and applications. *World Journal of Radiology*, *8*(9), 785–798. https://doi.org/10.4329/wjr.v8.i9.785

Bammer, R., Holdsworth, S. J., Veldhuis, W. B., & Skare, S. T. (2009). New methods in diffusion-weighted and diffusion tensor imaging. *Magnetic Resonance Imaging Clinics of North America*, *17*(2), 175–204. https://doi.org/10.1016/j.mric.2009.01.011

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*(1), 259–267. https://doi.org/10.1016/S0006-3495(94)80775-1

Bastiani, M., Cottaar, M., Dikranian, K., Ghosh, A., Zhang, H., Alexander, D. C., … Sotiropoulos, S. N. (2017). Improved tractography using asymmetric fibre orientation distributions. *NeuroImage*, *158*, 205–218. https://doi.org/10.1016/j.neuroimage.2017.06.050

Bastiani, M., Cottaar, M., Fitzgibbon, S. P., Suri, S., Alfaro-Almagro, F., Sotiropoulos, S. N., … Andersson, J. L. R. (2019). Automated quality control for within and between studies diffusion MRI data using a non-parametric framework for movement and distortion correction. *NeuroImage*, *184*, 801–812. https://doi.org/10.1016/j.neuroimage.2018.09.073

Bastin, M. E., Armitage, P. A., & Marshall, I. (1998). A theoretical study of the effect of experimental noise on the measurement of anisotropy in diffusion imaging. *Magnetic Resonance Imaging*, *16*(7), 773–785. https://doi.org/10.1016/S0730-725x(98)00098-8

Bates, D., Machler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baum, G. L., Roalf, D. R., Cook, P. A., Ciric, R., Rosen, A. F. G., Xia, C., … Satterthwaite, T. D. (2018). The impact of in-scanner head motion on structural connectivity derived from diffusion MRI. *NeuroImage*, *173*, 275–286. https://doi.org/10.1016/j.neuroimage.2018.02.041

Beaulieu, C. (2009). The biological basis of diffusion anisotropy. In H. Johansen-Berg & T. E. J. Behrens (Eds.), *Diffusion MRI: From Quantitative Measurement to In Vivo Neuroanatomy* (pp. 105–126). New York, NY: Academic Press. https://doi.org/10.1016/B978-0-12-374709-9.00006-7

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

Behrens, T. E., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., … Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, *50*(5), 1077–1088. https://doi.org/10.1002/mrm.10609

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, *29*(4), 1165–1188.

Boretius, S., Escher, A., Dallenga, T., Wrzos, C., Tammer, R., Bruck, W., … Stadelmann, C. (2012). Assessment of lesion pathology in a new animal model of MS by multiparametric MRI and DTI. *NeuroImage*, *59*(3), 2678–2688. https://doi.org/10.1016/j.neuroimage.2011.08.051

Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic resonance imaging: Physical principles and sequence design* (2nd ed.). New York, NY: WILEY Blackwell.

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. *Developmental Cognitive Neuroscience*, *32*, 43–54. https://doi.org/10.1016/j.dcn.2018.03.001

Chan, R. W., von Deuster, C., Giese, D., Stoeck, C. T., Harmer, J., Aitken, A. P., … Kozerke, S. (2014). Characterization and correction of eddy-current artifacts in unipolar and bipolar diffusion sequences

using magnetic field monitoring. *Journal of Magnetic Resonance*, *244*, 74–84. https://doi.org/10.1016/j.jmr.2014.04.018

Chanraud, S., Zahr, N., Sullivan, E. V., & Pfefferbaum, A. (2010). MR diffusion tensor imaging: A window into white matter integrity of the working brain. *Neuropsychology Review*, *20*(2), 209–225. https://doi.org/10.1007/s11065-010-9129-7

Chilla, G. S., Tan, C. H., Xu, C., & Poh, C. L. (2015). Diffusion weighted magnetic resonance imaging and its recent trend-a survey. *Quantitative Imaging in Medicine and Surgery*, *5*(3), 407–422. https://doi.org/10.3978/j.issn.2223-4292.2015.03.01

Dosenbach, N. U. F., Koller, J. M., Earl, E. A., Miranda-Dominguez, O., Klein, R. L., Van, A. N., … Fair, D. A. (2017). Real-time motion analytics during brain MRI improve data quality and reduce costs. *NeuroImage*, *161*, 80–93. https://doi.org/10.1016/j.neuroimage.2017.08.025

Einstein, A., Fürth, R., & Cowper, A. D. (1926). *Investigations on the theory of the Brownian movement*. London: Methuen & Co. Ltd.

Fortin, J. P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., … Hansen, K. D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology*, *15*(12), 503. https://doi.org/10.1186/s13059-014-0503-2

Fortin, J. P., Parker, D., Tunc, B., Watanabe, T., Elliott, M. A., Ruparel, K., … Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, *161*, 149–170. https://doi.org/10.1016/j.neuroimage.2017.08.047

Fortin, J. P., Sweeney, E. M., Muschelli, J., Crainiceanu, C. M., Shinohara, R. T., & Alzheimer's Disease Neuroimaging Initiative. (2016). Removing inter-subject technical variability in magnetic resonance imaging studies. *NeuroImage*, *132*, 198–212. https://doi.org/10.1016/j.neuroimage.2016.02.036

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*, 160044. https://doi.org/10.1038/sdata.2016.44

Greene, D. J., Koller, J. M., Hampton, J. M., Wesevich, V., Van, A. N., Nguyen, A. L., … Dosenbach, N. U. F. (2018). Behavioral interventions for reducing head motion during MRI scans in children. *NeuroImage*, *171*, 234–245. https://doi.org/10.1016/j.neuroimage.2018.01.023

Gruner, P., Vo, A., Ikuta, T., Mahon, K., Peters, B. D., Malhotra, A. K., … Szeszko, P. R. (2012). White matter abnormalities in pediatric obsessive-compulsive disorder. *Neuropsychopharmacology*, *37*(12), 2730–2739. https://doi.org/10.1038/npp.2012.138

Helmer, K. G., Chou, M. C., Preciado, R. I., Gimi, B., Rollins, N. K., Song, A., … Mori, S. (2016). Multi-site study of diffusion metric variability: Effects of site, vendor, field strength, and echo time on regions-of-interest and histogram-bin analyses. *Proceedings of SPIE The International Society for Optical Engineering*, *9788*, 97882U. https://doi.org/10.1117/12.2217445

Hrabe, J., Kaur, G., & Guilfoyle, D. N. (2007). Principles and limitations of NMR diffusion measurements. *Medical Physics*, *32*(1), 34–42. https://doi.org/10.4103/0971-6203.31148

Huo, J., Alger, J., Kim, H., Brown, M., Okada, K., Pope, W., & Goldin, J. (2016). Between-scanner and between-visit variation in Normal white matter apparent diffusion coefficient values in the setting of a multicenter clinical trial. *Clinical Neuroradiology*, *26*(4), 423–430. https://doi.org/10.1007/s00062-015-0381-3

Ismail, M. M., Keynton, R. S., Mostapha, M. M., ElTanboly, A. H., Casanova, M. F., Gimel'farb, G. L., & El-Baz, A. (2016). Studying autism Spectrum disorder with structural and diffusion magnetic resonance imaging: A survey. *Frontiers in Human Neuroscience*, *10*, 211. https://doi.org/10.3389/fnhum.2016.00211

Jayarajan, R. N., Venkatasubramanian, G., Viswanath, B., Janardhan Reddy, Y. C., Srinath, S., Vasudev, M. K., & Chandrashekar, C. R. (2012). White matter abnormalities in children and adolescents with obsessive-compulsive disorder: A diffusion tensor imaging study. *Depression and Anxiety*, *29*(9), 780–788. https://doi.org/10.1002/da.21890

Jbabdi, S., & Johansen-Berg, H. (2011). Tractography: Where do we go from here? *Brain Connectivity*, *1*(3), 169–183. https://doi.org/10.1089/brain.2011.0033

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156.

Jensen, J. H., & Helpern, J. A. (2010). MRI quantification of non-Gaussian water diffusion by kurtosis analysis. *NMR in Biomedicine*, *23*(7), 698–710. https://doi.org/10.1002/nbm.1518

Jeurissen, B., Leemans, A., Tournier, J. D., Jones, D. K., & Sijbers, J. (2013). Investigating the prevalence of complex fiber configurations in white matter tissue with diffusion magnetic resonance imaging. *Human Brain Mapping*, *34*(11), 2747–2766. https://doi.org/10.1002/hbm.22099

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118–127. https://doi.org/10.1093/biostatistics/kxj037

Jones, D. K., & Cercignani, M. (2010). Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine*, *23*(7), 803–820. https://doi.org/10.1002/nbm.1543

Jones, D. K., Knosche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–254. https://doi.org/10.1016/j.neuroimage.2012.06.081

Koirala, N., Anwar, A. R., Ciolac, D., Glaser, M., Pintea, B., Deuschl, G., … Groppa, S. (2019). Alterations in white matter network and microstructural integrity differentiate Parkinson's disease patients and healthy subjects. *Frontiers in Aging Neuroscience*, *11*, 191. https://doi.org/10.3389/fnagi.2019.00191

Koirala, N., Fleischer, V., Glaser, M., Zeuner, K. E., Deuschl, G., Volkmann, J., … Groppa, S. (2017). Frontal lobe connectivity and network community characteristics are associated with the outcome of subthalamic nucleus deep brain stimulation in patients with Parkinson's disease. *Brain Topography*, *31*, 311–321. https://doi.org/10.1007/s10548-017-0597-4

Koirala, N., Fleischer, V., Granert, O., Deuschl, G., Muthuraman, M., & Groppa, S. (2016). Network effects and pathways in deep brain stimulation in Parkinson's disease. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2016*, 5533–5536. https://doi.org/10.1109/embc.2016.7591980

Koirala, N., Perdue, M. V., Su, X., Grigorenko, E. L., & Landi, N. (2021). Neurite density and arborization is associated with reading skill and phonological processing in children. *NeuroImage*, *241*, 118426. https://doi.org/10.1016/j.neuroimage.2021.118426

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Landman, B. A., Farrell, J. A. D., Jones, C. K., Smith, S. A., Prince, J. L., & Mori, S. (2007). Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. *NeuroImage*, *36*(4), 1123–1138. https://doi.org/10.1016/j.neuroimage.2007.02.056

Le Bihan, D., Poupon, C., Amadon, A., & Lethimonnier, F. (2006). Artifacts and pitfalls in diffusion MRI. *Journal of Magnetic Resonance Imaging*, *24*(3), 478–488. https://doi.org/10.1002/jmri.20683

Leek, J. T., & Storey, J. D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genetics*, *3*(9), 1724–1735. https://doi.org/10.1371/journal.pgen.0030161

Lenth, R., & Love, J. (2018). lsmeans: Least-squares means. R package version 2.27-62.

Li, X., Yang, J., Gao, J., Luo, X., Zhou, Z., Hu, Y., … Wan, M. (2014). A robust post-processing workflow for datasets with motion artifacts in diffusion kurtosis imaging. *PLoS One*, *9*(4), e94592. https://doi.org/10.1371/journal.pone.0094592

Ling, J., Merideth, F., Caprihan, A., Pena, A., Teshiba, T., & Mayer, A. R. (2012). Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. *Human Brain Mapping*, *33*(1), 50–62. https://doi.org/10.1002/hbm.21192

Liu, S. Y., Thung, K. H., Lin, W. L., Yap, P. T., Shen, D. G., & Connectome, U. U. B. (2019). Multi-stage image quality assessment of diffusion MRI via semi-supervised nonlocal residual networks. In D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, et al. (Eds.), *Medical Image Computing and Computer Assisted Intervention. MICCAI 2019*. *Lecture Notes in Computer Science* (Vol. 11766, pp. 521–528). Cham: Springer. https://doi.org/10.1007/978-3-030-32248-9_58

Lu, H. Z., Jensen, J. H., Ramani, A., & Helpern, J. A. (2006). Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. *NMR in Biomedicine*, *19*(2), 236–247. https://doi.org/10.1002/nbm.1020

Maximov, I. I., Alnaes, D., & Westlye, L. T. (2019). Towards an optimised processing pipeline for diffusion magnetic resonance imaging data: Effects of artefact corrections on diffusion metrics and their age associations in UK biobank. *Human Brain Mapping*, *40*(14), 4146–4162. https://doi.org/10.1002/hbm.24691

Merboldt, K. D., Hanicke, W., & Frahm, J. (1985). Self-diffusion Nmr imaging using stimulated echoes. *Journal of Magnetic Resonance*, *64*(3), 479–486. https://doi.org/10.1016/0022-2364(85)90111-8

Mirzaalian, H., Ning, L., Savadjiev, P., Pasternak, O., Bouix, S., Michailovich, O., … Rathi, Y. (2016). Inter-site and inter-scanner diffusion MRI data harmonization. *NeuroImage*, *135*, 311–323. https://doi.org/10.1016/j.neuroimage.2016.04.041

Nana, R., Zhao, T., & Hu, X. (2008). Single-shot multiecho parallel echo-planar imaging (EPI) for diffusion tensor imaging (DTI) with improved signal-to-noise ratio (SNR) and reduced distortion. *Magnetic Resonance in Medicine*, *60*(6), 1512–1517. https://doi.org/10.1002/mrm.21770

Ning, L., Bonet-Carne, E., Grussu, F., Sepehrband, F., Kaden, E., Veraart, J., … Tax, C. M. W. (2020). Cross-scanner and cross-protocol multi-shell diffusion MRI data harmonization: Algorithms and results. *NeuroImage*, *221*, 117128. https://doi.org/10.1016/j.neuroimage.2020.117128

Nolte, U. G., Finsterbusch, J., & Frahm, J. (2000). Rapid isotropic diffusion mapping without susceptibility artifacts: Whole brain studies using diffusion-weighted single-shot STEAM MR imaging. *Magnetic Resonance in Medicine*, *44*(5), 731–736. https://doi.org/10.1002/1522-2594(200011)44:5<731::Aid-Mrm11>3.0.Co;2-1

Oguz, I., Farzinfar, M., Matsui, J., Budin, F., Liu, Z., Gerig, G., … Styner, M. (2014). DTIPrep: Quality control of diffusion-weighted images. *Frontiers in Neuroinformatics*, *8*, 4. https://doi.org/10.3389/fninf.2014.00004

Ozcan, A. (2010). A new model for diffusion weighted MRI: Complete Fourier direct MRI. *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, pp. 2710–2713.

Perrone, D., Aelterman, J., Pizurica, A., Jeurissen, B., Philips, W., & Leemans, A. (2015). The effect of Gibbs ringing artifacts on measures derived from diffusion MRI. *NeuroImage*, *120*, 441–455. https://doi.org/10.1016/j.neuroimage.2015.06.068

Pierpaoli, C. (2012). Artifacts in diffusion MRI. In D. K. Jones (Ed.), *Diffusion MRI: Theory, methods, and applications*. Oxford: Oxford Univeristy Press.

Pierpaoli, C., Jezzard, P., Basser, P. J., Barnett, A., & Di Chiro, G. (1996). Diffusion tensor MR imaging of the human brain. *Radiology*, *201*(3), 637–648. https://doi.org/10.1148/radiology.201.3.8939209

Pinto, M. S., Paolella, R., Billiet, T., Van Dyck, P., Guns, P. J., Jeurissen, B., … Sijbers, J. (2020). Harmonization of brain diffusion MRI: Concepts and methods. *Frontiers in Neuroscience*, *14*, 396. https://doi.org/10.3389/fnins.2020.00396

Polders, D. L., Leemans, A., Hendrikse, J., Donahue, M. J., Luijten, P. R., & Hoogduin, J. M. (2011). Signal to noise ratio and uncertainty in diffusion tensor imaging at 1.5, 3.0, and 7.0 tesla. *Journal of Magnetic Resonance Imaging*, *33*(6), 1456–1463. https://doi.org/10.1002/jmri.22554

Roalf, D. R., Quarmley, M., Elliott, M. A., Satterthwaite, T. D., Vandekar, S. N., Ruparel, K., … Gur, R. E. (2016). The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *NeuroImage*, *125*, 903–919. https://doi.org/10.1016/j.neuroimage.2015.10.068

Sasaki, M., Yamada, K., Watanabe, Y., Matsui, M., Ida, M., Fujiwara, S., … Investigators, A. J. (2008). Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: A multivendor multi-institutional comparison study. *Radiology*, *249*(2), 624–630. https://doi.org/10.1148/radiol.2492071681

Schmeel, F. C. (2019). Variability in quantitative diffusion-weighted MR imaging (DWI) across different scanners and imaging sites: Is there a potential consensus that can help reducing the limits of expected bias? *European Radiology*, *29*(5), 2243–2245. https://doi.org/10.1007/s00330-018-5866-4

Silk, T., Chen, J., Seal, M., & Vance, A. (2013). White matter abnormalities in pediatric obsessive-compulsive disorder. *Psychiatry Research*, *213*(2), 154–160. https://doi.org/10.1016/j.pscychresns.2013.04.003

Skare, S., Li, T. Q., Nordell, B., & Ingvar, M. (2000). Noise considerations in the determination of diffusion tensor anisotropy. *Magnetic Resonance Imaging*, *18*(6), 659–669. https://doi.org/10.1016/S0730-725x(00)00153-3

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

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*(Suppl 1), S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Soares, J. M., Marques, P., Alves, V., & Sousa, N. (2013). A Hitchhiker's guide to diffusion tensor imaging. *Frontiers in Neuroscience*, *7*, 31. https://doi.org/10.3389/fnins.2013.00031

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. *Developmental Cognitive Neuroscience*, *33*, 161–175. https://doi.org/10.1016/j.dcn.2017.12.002

te Grotenhuis, M., Pelzer, B., Eisinga, R., Nieuwenhuis, R., Schmidt-Catran, A., & Konig, R. (2017). When size matters: Advantages of weighted effect coding in observational studies. *International Journal of Public Health*, *62*(1), 163–167. https://doi.org/10.1007/s00038-016-0901-1

Theys, C., Wouters, J., & Ghesquiere, P. (2014). Diffusion tensor imaging and resting-state functional MRI-scanning in 5- and 6-year-old children: Training protocol and motion assessment. *PLoS One*, *9*(4), e94019. https://doi.org/10.1371/journal.pone.0094019

Tournier, J. D., Mori, S., & Leemans, A. (2011). Diffusion tensor imaging and beyond. *Magnetic Resonance in Medicine*, *65*(6), 1532–1556. https://doi.org/10.1002/mrm.22924

Travers, B. G., Adluru, N., Ennis, C., do Tromp, P. M., Destiche, D., Doran, S., … Alexander, A. L. (2012). Diffusion tensor imaging in autism spectrum disorder: A review. *Autism Research*, *5*(5), 289–313. https://doi.org/10.1002/aur.1243

Truong, T. K., Chen, N. K., & Song, A. W. (2011). Dynamic correction of artifacts due to susceptibility effects and time-varying eddy currents in diffusion tensor imaging. *NeuroImage*, *57*(4), 1343–1347. https://doi.org/10.1016/j.neuroimage.2011.06.008

Tuch, D. S., Reese, T. G., Wiegell, M. R., Makris, N., Belliveau, J. W., & Wedeen, V. J. (2002). High angular resolution diffusion imaging reveals

intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine*, *48*(4), 577–582. https://doi.org/10.1002/mrm.10268

Tuch, D. S., Reese, T. G., Wiegell, M. R., & Wedeen, V. J. (2003). Diffusion MRI of complex neural architecture. *Neuron*, *40*(5), 885–895. https://doi.org/10.1016/s0896-6273(03)00758-x

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., … Consortium, W. U.-M. H. (2012). The human Connectome project: A data acquisition perspective. *NeuroImage*, *62*(4), 2222–2231. https://doi.org/10.1016/j.neuroimage.2012.02.018

van Ewijk, H., Heslenfeld, D. J., Zwiers, M. P., Buitelaar, J. K., & Oosterlaan, J. (2012). Diffusion tensor imaging in attention deficit/hyperactivity disorder: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1093–1106. https://doi.org/10.1016/j.neubiorev.2012.01.003

Vandermosten, M., Boets, B., Wouters, J., & Ghesquiere, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience and Biobehavioral Reviews*, *36*(6), 1532–1552. https://doi.org/10.1016/j.neubiorev.2012.04.002

Veenith, T. V., Carter, E., Grossac, J., Newcombe, V. F. J., Outtrim, J. G., Lupson, V., … Coles, J. P. (2013). Inter subject variability and reproducibility of diffusion tensor imaging within and between different imaging sessions. *PLoS One*, *8*(6), e65941. https://doi.org/10.1371/journal.pone.0065941

Wang, Y., Mauer, M. V., Raney, T., Peysakhovich, B., Becker, B. L. C., Sliva, D. D., & Gaab, N. (2017). Development of tract-specific white matter pathways during early Reading development in at-risk children and typical controls. *Cerebral Cortex*, *27*(4), 2469–2485. https://doi.org/10.1093/cercor/bhw095

Warrington, S., Bryant, K. L., Khrapitchev, A. A., Sallet, J., Charquero-Ballester, M., Douaud, G., … Sotiropoulos, S. N. (2020). XTRACT—Standardised protocols for automated tractography in the human and macaque brain. *NeuroImage*, *217*, 116923. https://doi.org/10.1016/j.neuroimage.2020.116923

Webster, J. G., & Descoteaux, M. (2015). High Angular Resolution Diffusion Imaging (HARDI). In *Wiley Encyclopedia of Electrical and Electronics Engineering* (pp. 1–25). Hoboken, NJ: Wiley.

Wedeen, V. J., Hagmann, P., Tseng, W. Y., Reese, T. G., & Weisskoff, R. M. (2005). Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic Resonance in Medicine*, *54*(6), 1377–1386. https://doi.org/10.1002/mrm.20642

Wilkins, B., Lee, N., Gajawelli, N., Law, M., & Lepore, N. (2015). Fiber estimation and tractography in diffusion MRI: Development of simulated brain images and comparison of multi-fiber analysis methods at clinical b-values. *NeuroImage*, *109*, 341–356. https://doi.org/10.1016/j.neuroimage.2014.12.060

Wu, Z. M., Bralten, J., Cao, Q. J., Hoogman, M., Zwiers, M. P., An, L., … Wang, Y. F. (2017). White matter microstructural alterations in children with ADHD: Categorical and dimensional perspectives. *Neuropsychopharmacology*, *42*(2), 572–580. https://doi.org/10.1038/npp.2016.223

Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M., & Wandell, B. A. (2012). Development of white matter and reading skills. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(44), E3045–E3053. https://doi.org/10.1073/pnas.1206792109

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. https://doi.org/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. https://doi.org/10.1016/j.neuroimage.2012.03.072

How to cite this article: Koirala, N., Kleinman, D., Perdue, M. V., Su, X., Villa, M., Grigorenko, E. L., & Landi, N. (2022). Widespread effects of dMRI data quality on diffusion measures in children. *Human Brain Mapping*, *43*(4), 1326–1341. https://doi.org/10.1002/hbm.25724