Neuroradiology (2014) 56:497–510
DOI 10.1007/s00234-014-1342-2

FUNCTIONAL NEURORADIOLOGY

Reproducibility of diffusion tensor imaging in normal subjects: an evaluation of different gradient sampling schemes and registration algorithm

Xin Liu · Yong Yang · Jubao Sun · Gang Yu · Jin Xu · Chen Niu · Hongjun Tian · Pan Lin

Received: 16 December 2013 / Accepted: 13 February 2014 / Published online: 8 March 2014

Abstract
Introduction Diffusion tensor imaging (DTI) is very useful for investigating white matter integrity in ageing and neurological disorders; thus, evaluating its reproducibility under different acquisition protocols and analysis methods may assist in the design of clinical studies.
Methods To measure the reproducibility of DTI in normal subjects, this study include (1) depicting the reproducibility of DTI measurements in commonly used regions-of-interest analysis by intraclass correlation coefficient (ICC) and coefficient of variation (CV), (2) evaluating and comparing inter and intrasession test-retest reproducibility, and (3) illustrating the effect of the number of diffusion-encoding directions (NDED) and registration algorithms on measurement reproducibility.
Results DTI measurements exhibit high reproducibility, with overall (430/480) ICC≥0.70, (478/480) within-subject CV (CV_{ws}) ≤10.00 % and between-subject CV (CV_{bs}) ranging from 1.32 to 13.63 %. Repeated measures ANOVAs and paired *t* tests were conducted to compare inter and intrasession reproducibility with different diffusion sampling schemes and registration algorithms. Our results also confirmed that increasing the NDED could improve the accuracy and reproducibility of DTI measurements. In addition, we compared reproducibility indices that were derived using different registration algorithms, and a tensor-based deformable registration yielded the most reproducible results. Finally, we found that increasing the NDED could reduce the difference between the reproducibility of measurement derived using different registration algorithms and between the reproducibility of intersession and intrasession.
Conclusion Our results suggest that the choice of DTI acquisition protocol and post-processing methods can influence the accurate estimation and reproducibility of DTI measurements and should be considered carefully for clinical applications.

Keywords Diffusion tensor imaging · Reproducibility · Diffusion-encoding directions · Registration

Xin Liu and Jubao Sun contributed equally to the manuscript.

Electronic supplementary material The online version of this article (doi:10.1007/s00234-014-1342-2) contains supplementary material, which is available to authorized users.

X. Liu · J. Xu · P. Lin (✉)
Key Laboratory of Biomedical Information Engineering of Education Ministry, Institute of Biomedical Engineering, Xi'an Jiaotong University, No. 28, Xianning West Road, Xi'an, Shaanxi 710049, People's Republic of China
e-mail: linpan@mail.xjtu.edu.cn

J. Sun
MRI Center, The First Affiliated Hospital of Henan University of Science and Technology, Luoyang, Henan, People's Republic of China

Y. Yang
School of Information Technology, Jiangxi University of Finance and Economics, Nanchang, People's Republic of China

G. Yu
School of Info-physics and Geomatics Engineering, Central South University, Changsha, People's Republic of China

C. Niu
Department of Medical Imaging, The First Affiliated Hospital of Medical College, Xi'an Jiaotong University, Xi'an, People's Republic of China

H. Tian
Nanjing Fullshare Superconducting Technology Company Limited, Nanjing, People's Republic of China

Introduction

Magnetic resonance (MR) diffusion tensor imaging (DTI) is sensitive to water diffusion in tissues and is widely used to

noninvasively investigate white matter architecture [1]. DTI has become a popular tool for evaluating white matter integrity and structural connectivity. Diffusion tensor is a mathematical description of the diffusion process that yields various quantitative measurements, such as fractional anisotropy (FA) and mean diffusivity (MD), which reveal valuable information regarding the tissue microstructure. Because DTI has shown significant potential in addressing clinical and neurological issues, it is paramount to fully characterize its reproducibility for this method to become a standard modality in clinics.

DTI has become one of the most popular MRI techniques for the investigation of white matter architecture in normal and diseased brains. Prominent scalar quantities, including MD and FA, have been widely used in studies of neurodevelopment, reporting decreased MD and increased FA with brain maturation [2, 3], as well as in studies of neurological disorders, including amyotrophic lateral sclerosis [4–6], multiple sclerosis [7, 8], and stroke [9, 10]. Axial and radial diffusivities ($\lambda_{\parallel}$ and $\lambda_{\perp}$), in contrast, have also been proven to be useful in further obtaining white matter micro architecture. However, the use of axial and radial diffusivities to investigate pathology remains controversial, particularly in areas with crossing fibers [11]. Several studies of neurological diseases have employed axial and radial diffusivities as biomarkers, in which the alteration of these measurements indicate the impairment of white matter integrity [12–16]. However, previous studies of DTI reproducibility primarily focused on FA and MD, and limited data and studies have been reported regarding axial and radial diffusivities. Characterizing the reproducibility of axial and radial diffusivities is useful for investigating the underlying white matter architecture and integrity of the brain.

Previous studies have reported some factors that influence the reproducibility of DTI measurements, which includes the choice of acquisition schemes (i.e., diffusion-weighting schemes, number of scan repetition, etc.) and the post-processing pipeline (i.e., image registration, region-of-interest (ROI) drawing, fiber-tracking algorithms, etc.). Other uncontrollable factors include subject motion, physiological change, and MRI signal variation. The latter two factors are assumed to contribute higher variability over a longer period, which causes worse intersession reproducibility than that for intrasession. Wang [17] reported better intrasession reproducibility than that of intersession for DTI measurements in tractography and suggested that physiological change and MRI signal variation are major error sources in DTI tractography. Taken together, it is possible to determine major error sources that contribute to DTI measurement variability in ROI analysis by assessing and comparing inter and intrasession reproducibility.

Furthermore, the choice of optimal acquisition schemes is a controversial issue regarding the number of diffusion-encoding directions (NDED). To increase signal-to-noise ratio (SNR), some studies acquire repeated scans of the same set of diffusion-weighting directions or use more than the minimum six directions to increase the angular resolution of a DTI dataset [18–20]. Many simulation and experimental studies suggest that the latter is more preferable [17, 19]. However, the effect of the NDED on the accuracy and reproducibility of DTI measurements is rather controversial. One simulation study suggested that at least 20 unique diffusion-encoding directions are required for the robust estimation of anisotropy, and 30 unique diffusion-encoding directions are necessary for the robust estimation of tensor orientation [21]. In contrast, another study showed that with well-balanced sampling orientation, there is no advantage to increase the NDED from the minimal six directions [22] and that differences in DTI contrast due to different acquisition schemes are relatively smaller than intrasession variability [18]. In the present paper, we aim to investigate whether the influence of NDED is significant compared with other factors that contribute to measurement reproducibility and whether it is necessary to increase the NDED to achieve better data quality at the cost of a prolonged acquisition time.

The accurate analysis of DTI reproducibility is also dependent on analysis methodology. Several methods have been adopted to quantitatively analyze DTI measurements, including manually defined ROIs, histograms of the entire brain white matter [23], fiber tracking [24], and tract-based spatial statistics (TBSS) [25]. Voxel-wise and histogram analysis of manually defined ROIs are most commonly used. For this method, it is crucial to establish an accurate registration between images that are obtained in each session. The complexity in the registration of DTI images compared with scalar images lies in several aspects [26, 27]. To ensure the consistency of tensor orientation with anatomy, reorientation is often necessary after image transformation. In addition, similarity measurements that reflect all aspects of tensors, including size, shape, and orientation are not as easy to define as in scalar images. For simplicity, traditional DTI image registration is performed on scalar images (FA image, for example) and discards the orientation component [28–30]. Many studies have confirmed that such tensor-based registration algorithms were superior to traditional scalar map-based registration in their application [31, 32]. However, due to their complexity, these algorithms are not as widely used as traditional registration algorithms. Vollmar [33], who evaluated three different scalar image-based registration algorithms, suggested that the registration procedure could affect the reproducibility and sensitivity of the measurement. To the best of our knowledge, no studies have investigated whether tensor-based registration could improve DTI measurement reproducibility. In the present study, we comprehensively investigated four different registration schemes, including one tensor-based algorithm, and determined whether these schemes yield different reproducibility assessment results.

The goal of the present study was to characterize the test-retest reliability and between-subjects variability of DTI measurements using intraclass correlation coefficient (ICC) and coefficient of variation (CV) indices. The inter and intrasession reproducibility of FA, MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ was assessed to determine the main factors that contribute to the variability of DTI measurements. The effects of the NDED and image registration algorithms on DTI reproducibility were evaluated to determine the optimal acquisition scheme and post-processing procedures for clinical study.

Materials and methods

Subjects and imaging sessions

Nine healthy subjects (age range 20–40 years old, five males and four females) participated in this study. All subjects were fully informed of the nature of the study, and all subjects gave their consent regarding participation; the local ethical committee of our university institutional review board for clinical research approved all procedures. Subjects were scanned in two data sessions on separate days within a week. During the first session, two repeated scans were obtained; each scan was composed of three DTI datasets with 6, 15, and 32 isotropically distributed gradient sampling directions (called DTI_6A, DTI_6B, DTI_15A, DTI_15B, DTI_32A, and DTI_32B). During the second session, one such scan was obtained (called DTI_6C, DTI_15C, and DTI_32C).

Data acquisition

Data acquisition was performed using a 3.0T Philips Achieva scanner with a multiple channel SENSE head coil. Diffusion-weighted images were acquired using a single-shot EPI sequence (TE/TR=73.05/8,407.35 ms, 80 axial slices with slice thickness=2 mm), the imaging matrix was 112×112×80, with a 2-mm isotropic voxel size, FOV=224×224 mm. Diffusion weighting was applied using the b value=1000s/mm^2, with evenly distributed gradient sampling directions as described above. One additional b=0 image was acquired for each DTI datasets. Each DTI dataset was acquired for one time. The acquisition time for 6, 15, and 32 NDED sequence was 1:30, 3:03, and 6:01 min, respectively. A T1-weighted image was obtained for each subject using a 3D sequence. The imaging matrix was 256×256×176, with voxel size=1 mm^3, FOV=256×256 mm.

Image preprocessing and tenor fitting

DTI volumes were corrected for eddy current distortions by affine registration to the first non-diffusion-weighted volumes using FMRIB's linear image registration tool (FLIRT, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki) and by skull stripping using a BET (brain extraction tool in FMRIB Software Library). Diffusion tensors were fitted in each voxel using FDT (FMRIB's diffusion toolbox), and FA, MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ maps were calculated for individual sessions.

Measurement map registration

We employed four different registration schemes to assess reproducibility. For image-based registration, FA maps were selected as feature images and were registered to the template; the derived deformation field was applied to corresponding MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ maps to realign the maps to the template space.

1. Linear transformation was performed to register FA maps that were derived from three repeated scans of each subject to the FMRIB58_FA template using FMRIB's linear image registration tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FLIRT). The registered FA maps were used to calculate the mean FA map for each subject, which was then used as the individual template. The linear registration was repeated to register FA maps that were obtained from three scans to the individual template, and transformation matrixes were saved and applied to corresponding MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ maps. In this method, measurement maps were aligned to the FMRIB58_FA template while preserving each subject's individual anatomy (called FLIRTreg in this paper).
2. Deformable registration was performed to register each subject's FA maps, which were derived from three repeated scans, to the FMRIB58_FA template using FMRIB's non-linear image registration tool (FNIRT, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT), and the deformation field was applied to the corresponding MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ maps. ROIs were defined on the FMRIB58_FA template and were applied to individual measurement maps (called FNIRTreg in this paper).
3. Tensor-based deformable registration was implemented using DTI-TK (http://dti-tk.sourceforge.net/). The registration pipeline includes bootstrapping the initial DTI template from the input DTI volumes, performing deformable alignment following affine alignment with template refinement. The registered tensor data were used to generate measurement maps of three scans for each subject, and ROIs were drawn on FA maps that were derived from the template tensor (called DTITKreg in this paper).
4. Diffusion-weighted images from each scan within a single subject were registered using linear transformation (FMRIB's linear image registration tool). This registration was performed by selecting the b0 image from the first scan of each subject as the target image to which b0 images from the other two scans were registered. The

transformation matrix was then applied to the rest of the DWIs within the same dataset, and the gradient direction table was modified accordingly (called DWIreg in this paper).

Among the four registration algorithms, DWIreg and FLIRTreg utilized a linear transformation model, and ROIs were drawn in each subject's own diffusion space; FNIRTreg and DTITKreg utilized a deformable model, and ROIs were drawn in the template. DTITKreg utilized the full tensor information and was expected to produce the best results.

ROI drawing

The major white matter area in the corpus callosum and in the internal capsule was chosen as the region of interest. The ROI drawing was performed using the TrackVis software (http://trackvis.org/) by the same author under the inspection of an experienced radiologist. Restrictions for ROI drawing are described as follows and are shown in Fig. 1. The values of FA, MD, $\lambda_{\|}$, and $\lambda_{\perp}$ (mm^2/s) were extracted within each ROI.

The body of the corpus callosum (CC)	The entire area of the corpus callosum on the midsagittal plane was selected as the ROI.
The genu of the corpus callosum (GCC)	The GCC is anatomically defined as the anterior bulbous area of the corpus callosum. The GCC was drawn on consecutive sagittal planes to cover the region of the GCC.
The splenium of the corpus callosum (SCC)	The SCC was drawn on consecutive sagittal planes to cover the region of the SCC.
The posterior limb of the internal capsule (PIC)	The PIC was drawn on three consecutive slices on which the diameter of the PIC was the largest compared with adjacent slices.
The anterior limb of the internal capsule (AIC)	The AIC was drawn on the same slice as the PIC.

Statistical analysis

The reproducibility of DTI measurements was quantified using a two-way random model ICC with absolute agreement. The variability of each DTI measurements was quantified using CV, which is defined as the percentage of standard deviation of a set of measurements by its mean. Measurements with ICC≤0.40 were considered poor, whereas ICC≥0.80 and within-subject CV (CV_{ws}) ≤10 % were considered excellent [34, 35]. The intrasession reproducibility was estimated using contrasting repeated scans in the first session (i.e., DTI_6A and DTI_6B, etc.), whereas intersession reproducibility was estimated by contrasting the first scan that was obtained in each imaging session (i.e., DTI_6A and DTI_6C, etc.). A repeated-measure ANOVA and paired *t* tests were conducted on ICC and CV to determine the effects of the NDED and registration algorithms on the reliability of measurements.

Results

Tables S1, S2, S3, and S4 containing the ICC and CV results for FA, MD, $\lambda_{\|}$, and $\lambda_{\perp}$ in each registration algorithm and sequence are in the supplementary materials. We evaluated how the DTI acquisition protocol and the post-processing pipeline, in particular, the NDED and registration methods, could affect the estimate of DTI measurements and their reproducibility. Parametric tests were conducted on group mean values of measurements that were derived from ROIs as well as on ICC and CV indices. Cases with both intra and intersession ICC≥0.80 and CV_{ws}≤10.00 % are considered highly reproducible and are indicated in bold font in Tables S1, S2, S3, and S4. Figure 2 shows the mean values of DTI measurements that were observed in each sequence using FNIRTreg.

Study 1: the reproducibility of DTI measurement in different ROIs

In this study, the observed DTI measurements showed high reproducibility, with ICC ranging from 0.25 to 1.00 (90 % of ICC were higher than 0.70), CV_{ws} ranging from 0.30 to

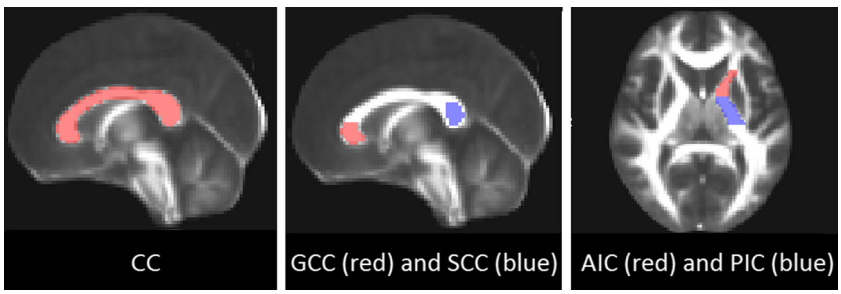

Fig. 1 Region-of-interest (ROI) drawing that is displayed on the images. Abbreviations: *CC*, corpus callosum; *GCC*, genu of the corpus callosum; *SCC*, splenium of the corpus callosum; *PIC*, posterior limb of the internal capsule; *AIC*, anterior limb of the internal capsule

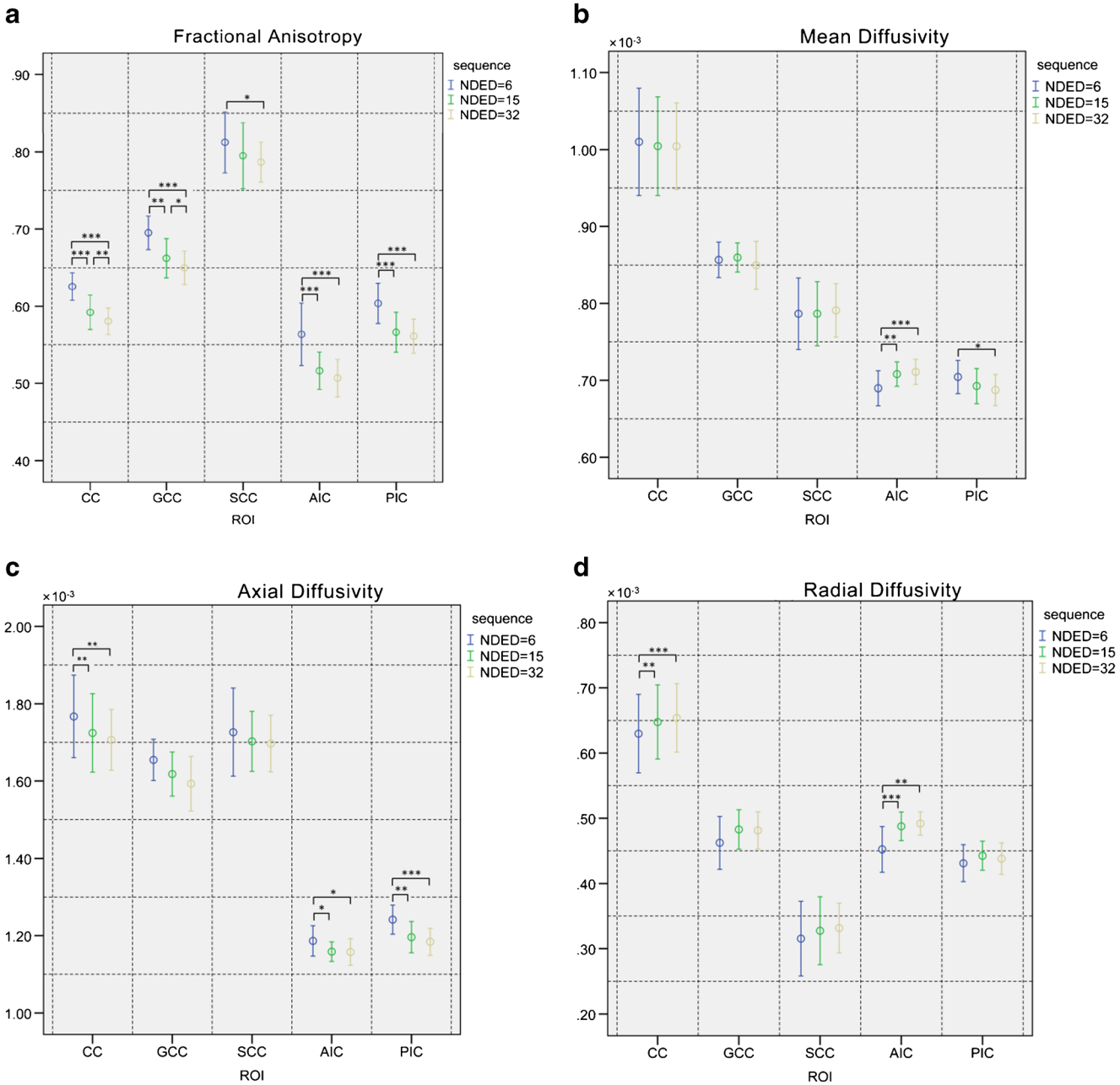

Fig. 2 (**a**) FA, (**b**) MD (mm^2/s), (**c**) $\lambda_{\parallel}$ (mm^2/s), and (**d**) $\lambda_{\perp}$ (mm^2/s) that were derived using sequences with 6, 15 and 32 NDED (average value and standard deviation from nine subjects). For simplicity, we only show results that were achieved using the FNIRTreg registration algorithm. Paired *t* test results comparing different sequences are shown in the figure. *$p<0.05$, **$p<0.01$, ***$p<0.001$. For abbreviations, see Fig. 1

13.01 % (97 % of CV_{ws} were lower than 5.00 %), and between-subject CV (CV_{bs}) ranging from 1.32 to 13.63 % (94 % of CV_{bs} were lower than 10.00 %). Across different DTI sequences and registration algorithms, measurements in CC showed higher consistency, with the highest number of ICC≥0.8 (For measurements that were derived using FNIRTreg, all ICC≥0.80 except ICC=0.69 for FA in six NDED sequences, as indicated in bold in Table S2) than those values in the other ROIs, whereas SCC showed poor consistency with five ICC≤0.40, followed by AIC with two ICC≤0.40. Measurements in SCC showed higher test-retest variability (For FNIRTreg, the mean CV_{ws}=2.77 % for SCC, whereas the mean CV_{ws}=1.06–1.34 % for the other four ROIs) with two CV_{ws}≥10 %.

Study 2: effects of the NDED on descriptive statistics

We first evaluated whether the choice of the NDED has an effect on the absolute value of derived DTI measurements. Figure 2 represents the average DTI measurements within

each ROI that was derived from the first scan using FNIRTreg, which is the image-based deformable registration that is most commonly used in the literature [36, 33]. We conducted a two-way ANOVA, which considers the factors of different sequences and ROIs, and found a significant impact of the NDED on derived FA, $\lambda_{\|}$, and $\lambda_{\perp}$ values (main factor of sequence, F=30.174; p<0.001 for FA; F=5.798, p=0.004 for $\lambda_{\|}$; F=3.840, p=0.024 for $\lambda_{\perp}$). A paired t test showed the consistent difference of FA between sequences across ROIs, as shown in Fig. 2 (diffusivity: mm^2/s).

Study 3: intersession and intrasession effects on reproducibility

We conducted paired t tests to compare intersession and intrasession ICC and CV_{ws} of each measurement across five ROIs (see Figs. 3 and 4). In six NDED sequences, FA showed significantly higher intersession CV_{ws} than that of intrasession using DWIreg (paired $t(4)$=3.178, p=0.034), and $\lambda_{\|}$ showed significantly higher intersession CV_{ws} (paired $t(4)$=3.518, p=0.024) and lower intersession ICC (paired $t(4)$=2.882, p=0.045) than that of intrasession using FLIRTreg. In 15 NDED sequences, FA showed significantly higher intersession CV_{ws} than that of intrasession using FLIRTreg (paired $t(4)$=2.777, p=0.05) and DWIreg (paired $t(4)$=4.508, p=0.011) and significantly lower intersession ICC than that of intrasession using DWIreg (paired $t(4)$=3.268, p=0.031) and DTITKreg (paired $t(4)$=6.000, p=0.004); $\lambda_{\|}$ showed significantly higher intersession CV_{ws} than that of intrasession using FLIRTreg (paired $t(4)$=3.957, p=0.017); and MD showed significantly lower intersession ICC than that of intrasession using FLIRTreg (paired $t(4)$=3.015, p=0.039) and DWIreg (paired $t(4)$=3.586, p=0.023). In 32 NDED sequences, FA showed significantly higher intersession CV_{ws} (paired $t(4)$=17.571, p<0.001) and lower intersession ICC than that of intrasession (paired $t(4)$=4.134, p=0.014) using DWIreg.

Study 4: factors influencing reproducibility–NDED and DTI measurements

In the present study, we observed a significant effect of DTI measurements and the NDED on reproducibility. In total, 80 sets of cases were acquired for each sequence (four measurement in five ROIs that were derived using four registration algorithms); among which, 39 were highly reproducible (both inter and intrasession ICC≥0.80 and CV_{ws}≤10 %) for sequences with 6 NDED, 55 for sequences with 15 NDED, and 61 for sequences with 32 NDED. In total, 60 sets of cases were acquired for each measurement (each measurement in five ROIs across three sequences that were derived using four registration algorithms); MD and $\lambda_{\perp}$ yield more reproducible cases than FA and $\lambda_{\|}$ do (47 for MD, 42 for $\lambda_{\perp}$, 35 for $\lambda_{\|}$, and 31 for FA). Although $\lambda_{\perp}$ appears to be reproducible, this measurement suffers from higher test-retest and intersubject variability (the mean CV_{ws}=2.59 % and the mean CV_{bs}= 7.73 % for $\lambda_{\perp}$, whereas the mean CV_{ws}=1.02–1.51 % and the mean CV_{bs}=3.68–4.20 % for other three measurements).

The effects of the NDED on DTI measurement reproducibility were further tested using repeated-measure ANOVA, as shown in Table 1. We also included factors of ROIs, measurements, and registration algorithms (and their second order interactions) in the model. The NDED have a significant impact on measurement reproducibility and on CV_{bs}. An interaction was observed between NDED and DTI measurements on ICC and CV, such that the effect of NDED was not as prominent on $\lambda_{\|}$ as on the other three measurements. For simplicity, we only report paired t test results on reproducibility indices that were derived using FNIRTreg, which is the most commonly used registration strategy. We found significantly lower intersession ICC for MD and $\lambda_{\perp}$, higher inter and intrasession CV_{ws} for FA, and higher CV_{bs} for $\lambda_{\perp}$ that was derived using a sequence with 6 NDED than that derived using a sequence with 32 NDED (p<0.05). We also observed significantly lower inter and intrasession ICC for MD and higher CV_{bs} for $\lambda_{\perp}$ that was derived using a sequence with 6 NDED than that derived using a sequence with 15 NDED (p<0.05). A sequence with 15 NDED only derived a higher intrasession CV_{ws} for FA when compared with a sequence with 32 NDED (p<0.05).

Study 5: factors influencing reproducibility–The image registration algorithm

Most CV and ICC that were derived from the four registration algorithms are within acceptable levels. We acquired 60 sets of cases using each algorithm (four measurements in five ROIs across three sequences), with 44 highly reproducible cases (both inter- and intrasession ICC≥0.80 and CV_{ws}≤10 %) that were observed for FLIRTreg, 34 for FNIRTreg, 25 for DWIreg, and 52 for DTITKreg.

The effects of registration algorithms on reproducibility were tested using a repeated-measure ANOVA, as shown in Table 2. We also included factors of the NDED, DTI measures, and ROIs (and their second order interactions) in our model. We observed a significant effect of the image registration algorithm on ICC and CV. An interaction was observed between the registration algorithm and DTI measurements. An interaction was also observed between the registration algorithm and NDED. For compactness, Figs. 5, 6, and 7 only

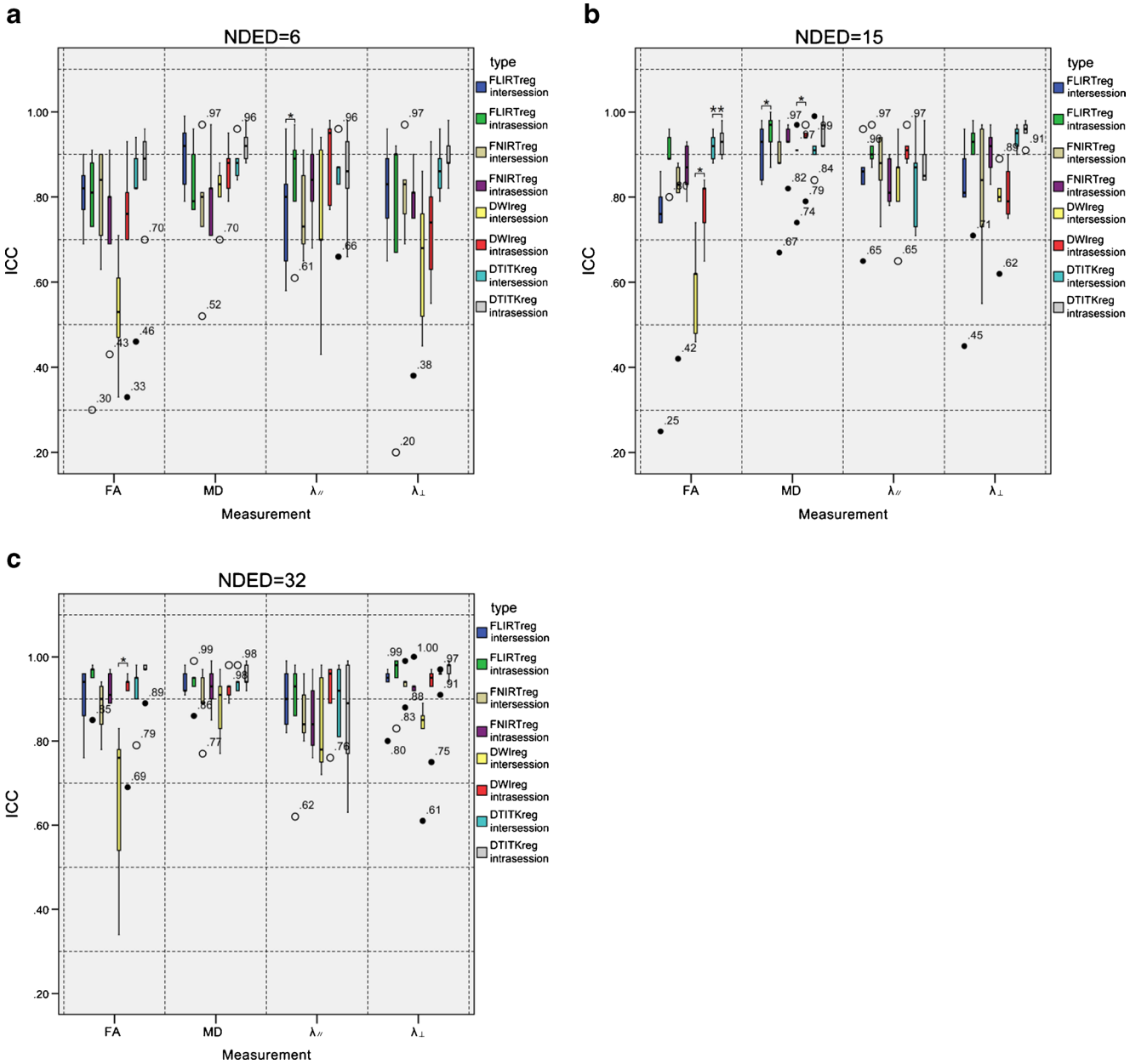

Fig. 3 Box plot contrasting inter and intrasession ICC of measurements that were derived using (**a**) a sequence with 6 NDED (**b**) a sequence with 15 NDED, and (**c**) a sequence with 32 NDED. Paired *t* test results comparing inter and intrasession ICC are shown in the figure. *$p<0.05$, **$p<0.01$, ***$p<0.001$. Abbreviations: NDED, number of diffusion encoding directions; ICC, intraclass correlation coefficient; CV, coefficient of variation; FA, fractional anisotropy, MD, mean diffusivity; $\lambda_{\parallel}$, axial diffusivity; $\lambda_{\perp}$, radial diffusivities

show intrasession ICC and CV that were derived using different registration algorithms and paired *t* test results.

Discussion

The evaluation of the reproducibility of DTI quantitative measurements is essential for clinical longitudinal studies. In the present study, we primarily focused on FA, MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$, which were derived using a ROI approach, and investigated how the acquisition protocol and the registration algorithm could affect the reproducibility and bias of DTI measurements.

The reproducibility of DTI measurements

In the present study, we evaluated five ROIs that are broadly used in studies of neurodevelopment and white matter

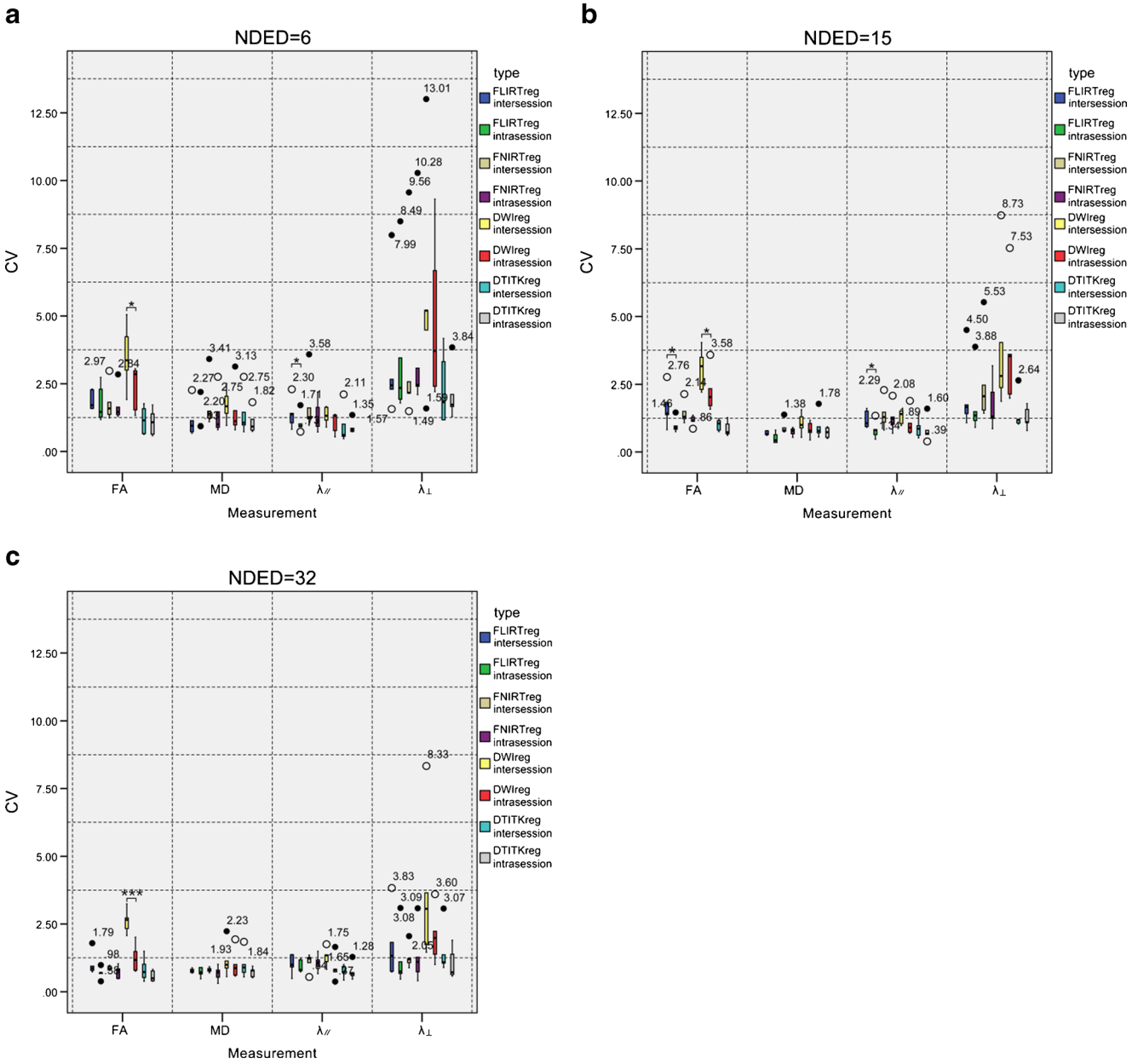

Fig. 4 Box plot contrasting inter and intrasession CV_{ws} of measurements that were derived using (**a**) a sequence with 6 NDED, (**b**) a sequence with 15 NDED, and (**c**) a sequence with 32 NDED. Paired *t* test results comparing inter and intrasession ICC are shown in the figure. $*p<0.05$, $**p<0.01$, $***p<0.001$. Abbreviations: NDED, number of diffusion encoding directions; ICC, intraclass correlation coefficient; CV, coefficient of variation; FA, fractional anisotropy, MD, mean diffusivity; $\lambda_{\parallel}$, axial diffusivity; $\lambda_{\perp}$, radial diffusivities

Table 1 Results of an ANOVA test of the effect of the NDED on reproducibility

Main factor	ICC		CV_{ws}		CV_{bs}
	Intrasession	Intersession	Intrasession	Intersession	
NDED	$F=108.373$, $p<0.001$	$F=37.812$, $p<0.001$	$F=110.481$, $p<0.001$	$F=92.861$, $p<0.001$	$F=26.541$, $p<0.001$
NDED measurement[a]	$F=11.489$, $p<0.001$	nonsignificant	$F=24.973$, $p<0.001$	$F=14.587$, $p<0.001$	$F=6.508$, $p<0.001$

NDED number of diffusion-encoding directions

[a] The interaction effect between NDED and measurement type

Table 2 Results of an ANOVA test of the effect of the registration algorithm on reproducibility

Main factor	ICC		CV_{ws}		CV_{bs}
	Intrasession	Intersession	Intrasession	Intersession	
Registration	F=18.920, p<0.001	F=39.924, p<0.001	F=48.327, p<0.001	F=125.030, p<0.001	F=53.627, p<0.001
Registration measurement[a]	F=6.862, p<0.001	F=7.905, p<0.001	F=11.920, p<0.001	F=29.140, p<0.001	F=12.168, p<0.001
Registration NDED[b]	F=5.452, p<0.001	F=2.580, p=0.025	F=4.299, p=0.003	F=3.593, p=0.009	F=2.415, p=0.035

[a] The interaction effect between registration algorithm and measurement type
[b] The interaction effect between registration algorithm and NDED
NDED number of diffusion encoding directions

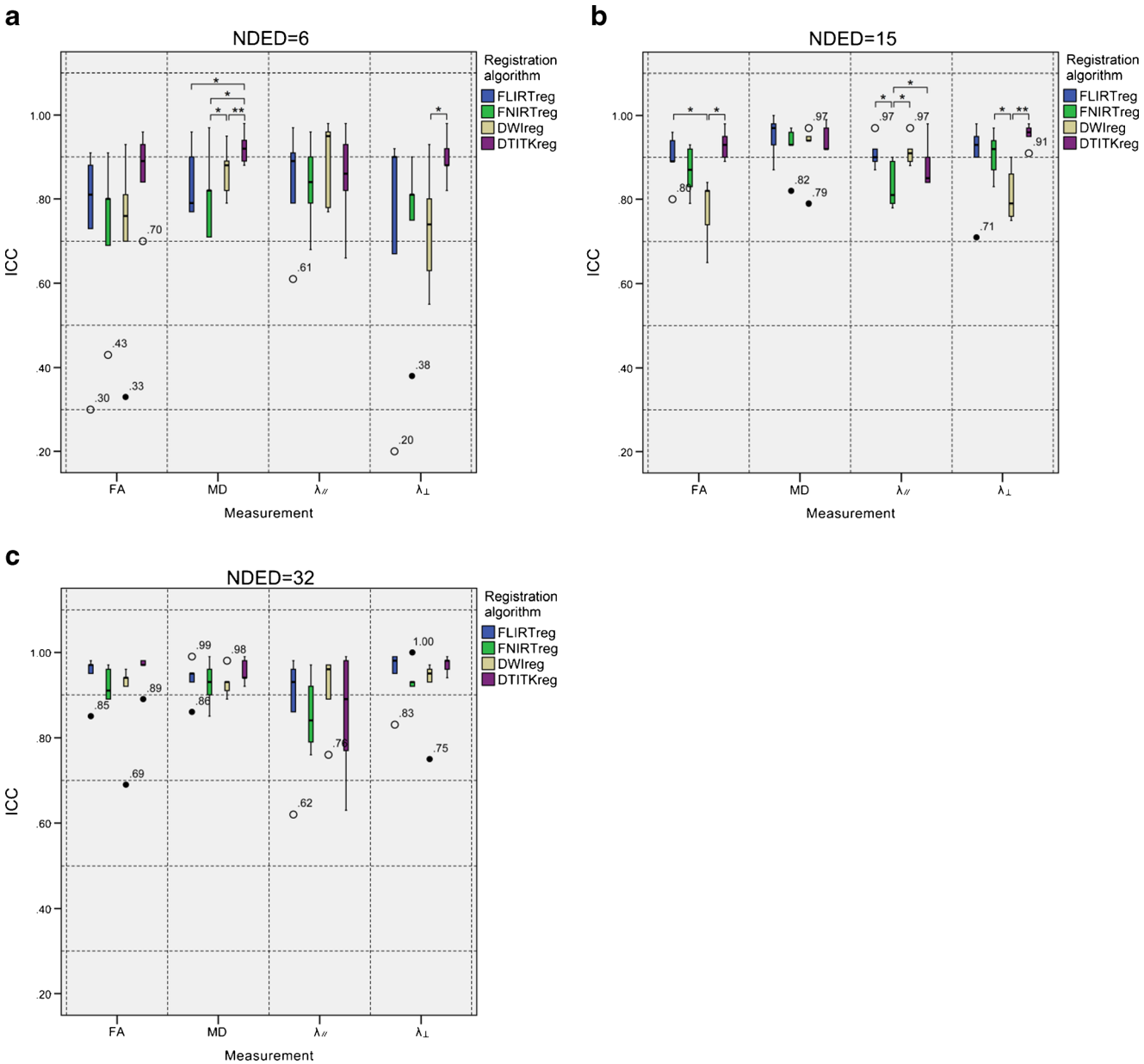

Fig. 5 Box plot depicting intrasession ICC of measurements that were achieved with different registration algorithms when using (**a**) a sequence with 6 NDED, (**b**) a sequence with 15 NDED, and (**c**) a sequence with 32 NDED. Paired *t* test results comparing ICC that were achieved with different registration algorithms are shown in the figure. *p<0.05, **p<0.01, ***p<0.001. Abbreviations: NDED, number of diffusion encoding directions; ICC, intraclass correlation coefficient; CV, coefficient of variation; FA, fractional anisotropy, MD, mean diffusivity; $\lambda_{\parallel}$, axial diffusivity; $\lambda_{\perp}$, radial diffusivities

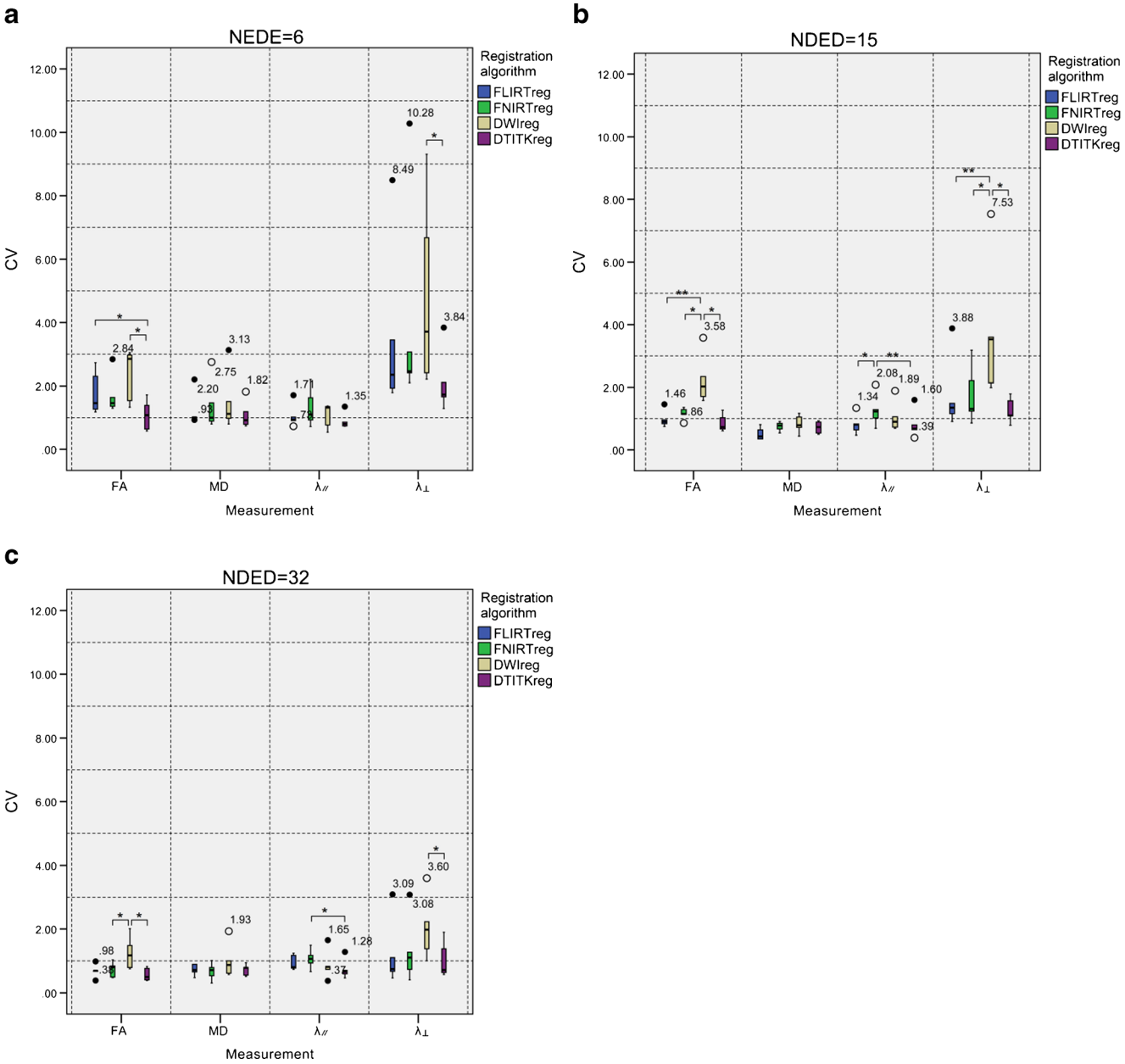

Fig. 6 Box plot depicting intrasession CV_{ws} of measurements that were achieved with different registration algorithms when using (**a**) a sequence with 6 NDED, (**b**) a sequence with 15 NDED, and (**c**) a sequence with 32 NDED. Paired *t* test results comparing CV_{ws} that were achieved with different registration algorithms are shown in the figure. *$p<0.05$, **$p<0.01$, ***$p<0.001$. Abbreviations: NDED, number of diffusion encoding directions; ICC, intraclass correlation coefficient; CV, coefficient of variation; FA, fractional anisotropy, MD, mean diffusivity; $\lambda_{\parallel}$, axial diffusivity; $\lambda_{\perp}$, radial diffusivities

diseases [37–42]. Consistent with previous studies, the reproducibility is measurement and ROI specific [35, 33, 36]. As described in the "Results" section, we found that measurements produced by CC showed higher reproducibility than measurements produced by SCC and AIC. To compare the reproducibility of different measurements, we found that MD yielded the highest ICC across DTI sequences and registration algorithms (the mean ICC=0.90), followed in turn by $\lambda_{\perp}$ (the mean ICC=0.85), $\lambda_{\parallel}$ (the mean ICC=0.85), and FA (the mean ICC=0.81). Although $\lambda_{\perp}$ showed high consistency by yielding acceptable ICC, this value suffered from high test-retest variability by producing higher CV_{ws} than the other three measurements, which could substantially limit its use in detecting small changes within subject over time. $\lambda_{\perp}$ also showed higher between-subjects variability by producing higher CV_{bs} than the other three measurements, which could hamper its use in investigating group differences. The dependency of reproducibility on DTI measurements and on ROI must be considered in designing experiments, i.e., sample size. The present study contributes to this issue.

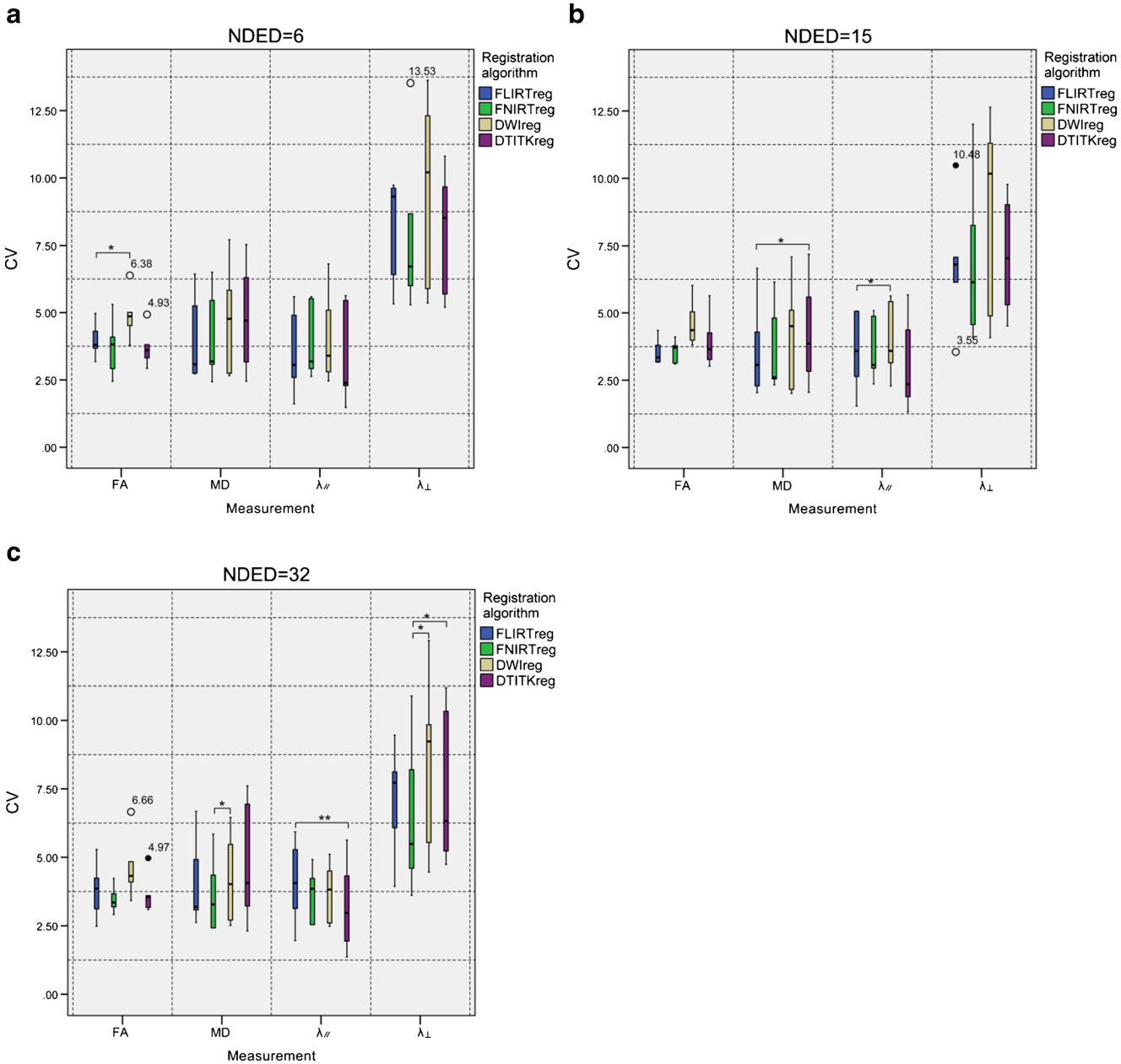

Fig. 7 Box plot depicting CV_{bs} of measurements that were achieved with different registration algorithms when using (**a**) a sequence with, 6 NDED, (**b**) a sequence with 15 NDED, and (**c**) a sequence with 32 NDED. Paired *t* test results comparing CV_{bs} that were achieved with different registration algorithms are shown in the figure. $*p<0.05$, $**p<0.01$, $***p<0.001$. Abbreviations: NDED, number of diffusion encoding directions; ICC, intraclass correlation coefficient; CV, coefficient of variation; FA, fractional anisotropy, MD, mean diffusivity; $\lambda_{||}$, axial diffusivity; $\lambda_{\perp}$, radial diffusivities

The intra and intersession reproducibility

We compared intrasession and intersession ICC and CV_{ws} to determine the major error source in DTI data. Wang [17] reported significantly lower intersession ICC and higher intersession CV_{ws} than those intrasession values for some DTI measurements in tractography, particularly for FA, which indicated that MRI signal variation and physiological changes over time were important error sources in tractography. In our study, we extended similar results to ROI analysis by comparing the inter and intrasession ICC and CV_{ws} across five ROIs using different DTI sequences and registration algorithms. The paired *t* test results showed significantly worse intersession reproducibility than that of intrasession, especially for FA, which could limit its use in longitudinal studies. This also confirmed that MRI signal variation and physiological change are major error sources in ROI analysis. In addition, the results suggest that increasing the NDED from 15 to 32 could reduce the

significance of inter and intrasession differences. In addition, inter and intrasession differences could become more prominent when using linear registration algorithms, with most sets of data showing significant inter and intrasession differences when using DWIreg and FLIRTreg registration algorithms.

The effect of the NDED on the estimation and reproducibility of measurements

Many studies suggested that increasing the number of diffusion encoding directions (NDED) leads to more robust tensor estimation. Simulation studies have illustrated that low SNR can cause an overestimation of the axial diffusivity and an underestimation of radial diffusivity [43], which results in overestimated FA, and a similar result is shown using experimental data [17, 44]. Heiervang [23] observed the overestimation of FA and the underestimation of MD in fewer NDED sequences. Ni and colleagues, in contrast, reported similar FA and MD values that were derived from sequences with different NDED, yet eigenvalues were affected by the choice of the NDED [19]. In the present study, we compared DTI measurements that were derived using sequences with 6, 15, and 32 NDED and found that increasing the NDED improves both the accurate estimation and the reproducibility of derived measurements. With higher angular resolution and as already shown by previous studies [45, 23], it is reasonable to consider measurements that are derived using 32 NDED sequences to be more accurate. We observed that acquiring data with less NDED introduces upward bias into FA and $\lambda_{\parallel}$ values, downward bias into $\lambda_{\perp}$, and mixed bias in the MD value either upward, downward, or unchanged. As shown in Fig. 2, the FA value that was obtained with different NDED sequences showed consistent differences across five ROIs, and no significant differences were found for MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ that were derived from 15 and 32 sequences. This result suggests that a sequence with 15 NDED is sufficient in providing an accurate estimate of MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$ values. In contrast, acquiring data using sequences with adequate NDED (at least 32) is necessary when using FA as a biomarker. In addition to the effect of NDED on the estimate of DTI measurement, we found that increasing NDED could also improve measurement reproducibility. According to the paired *t* test results that were mentioned in the previous section, the effect of NDED on reproducibility is also measurement specific, such that the effect of NDED was not as prominent on $\lambda_{\parallel}$ as on FA, MD, and $\lambda_{\perp}$.

The effect of the image registration algorithm on reproducibility

The registration of DTI images has an important application in assisting clinical studies to detect the variation in DTI measurements during brain development [46, 28] and pathological processes [31]. It is a common practice to perform deformable registration on DTI images of each subject to a common template. In this study, we evaluated two deformable registration algorithms, a scalar image-based registration FNIRTreg and a tensor-based registration DTITKreg. In general, DTITKreg yielded better results, with significant higher ICC (mean ICC=0.85 for FNIRTreg, mean ICC=0.90 for DTITKreg), lower CV_{ws} than FNIRTreg (mean CV_{ws}=1.52 for FNIRTreg, mean CV_{ws}=1.09 for DTITKreg), and similar CV_{bs} (mean CV_{bs}=4.59 for FNIRTreg, mean CV_{bs}=4.82 for DTITKreg). The comparison is shown in Figs. 5, 6, and 7. The advantage of DTITKreg over FNIRTreg is consistent with our previous assumption and likely results from the fact that DTITKreg makes full use of tensor information. According to Gee and Alexander [26], tensor-based registration algorithms, which consider the tensor shape, size, and orientation, could improve registration effectiveness by avoiding spurious kinks and whorls in homogeneous regions. In addition, we compared two image-based registrations with a different deformation model: the FNIRTreg utilizes a deformable model, whereas FLIRTreg uses a linear transformation. We found that $\lambda_{\parallel}$ that was obtained using 15 NDED sequences showed a higher intrasession ICC and CV_{ws} when using FLIRTreg compared with FNIRTreg, and MD obtained using 6 NDED sequences showed a lower intersession CV_{ws} when using FLIRTreg. Due to its convenience in group studies, image-based deformable registration algorithms, such as FNIRTreg, are most commonly used [33, 36]. However, in the present study, we found that image-based linear transformation is superior to its deformable counterpart in producing reproducible measurements. Similar to our result, Vollmar [33] observed a lower ICC of FA when using image-based deformable registration to a common template, which suggested that normalization to a template space could reduce intersubject variation, thus decreasing the ICC value. Furthermore, we compared two registration algorithms with linear transformation: FLIRTreg and DWIreg. FLIRTreg outperformed DWIreg most likely because the b0 image has rather smoothed appearance compared with the FA image and lacks the contrast to drive as good as spatial match as the FA image does.

Another major finding of the present study is that the effect of registration algorithms on reproducibility is sequence and measurement specific. As shown in Fig. 5, intrasession ICC of measurements that were obtained using sequences with 32 NDED showed no significant difference between registration algorithms, indicating that acquiring data with sufficient NDED could deemphasize the choice of image processing pipeline. In addition, as shown in Fig. 6, intrasession CV_{ws} of MD showed no significant difference between registration algorithms across three sequences, which indicated that the image processing strategy is less likely to affect MD reproducibility.

Summary of major findings

The present study provided a comprehensive analysis of the reproducibility of DTI measurements in ROI-based analysis. Our results suggest that the choice of DTI acquisition protocol and post-processing methods could influence the accurate estimation and reproducibility of DTI measurements, thus, should be considered carefully depending on the specific issue that the study aims to investigate. The major findings of the current study can be summarized as follows:

1. DTI measurements in ROI analysis exhibit high reproducibility. However, $\lambda_{\perp}$ suffered from high test-retest variability and between-subjects variability, which could substantially limit its use in detecting small changes within subject and investigating group differences.
2. Our results suggest that MRI signal variation and physiological change over time causes significantly worse intersession reproducibility than that of intrasession in ROI analysis, especially for FA.
3. Our results suggest that the currently commonly used image-based registration algorithm may not be optimal for group studies. We suggest that a tensor-based registration algorithm could improve the registration efficiency and thus provide more reproducible data in ROI analysis.
4. Our results confirmed that increasing the NDED could improve both the accurate estimation and reproducibility of DTI measurements. Beyond that, acquiring data with more NDED could reduce the difference between inter and intrasession reproducibility, thus promoting the use of DTI measurements in longitudinal studies. Acquiring data with more NDED may also deemphasize the choice of the image processing pipeline.

Acknowledgments This work was supported by the National Natural Science Foundation of China (nos. 60963012, 61262034, 11272242, and 31271061), the Fundamental Research Funds for the Central Universities, the Key Project of Chinese Ministry of Education (no. 211087), the Natural Science Foundation of Jiangxi Province (no. 20132BAB201025), the Young Scientist Foundation of Jiangxi Province (no. 20122BCB23017), the Science and Technology Research Project of the Education Department of Jiangxi Province (no. GJJ13302), the Central Universities of Central South University (2013zzts251), the Doctoral Fund of Ministry of Education of China (20120201120071), and the Fundamental Research Funds for the Central Universities of China.

Conflict of interest We declare that we have no conflict of interest.

References

1. Basser PJ, Mattiello J, LeBihan D (1994) MR diffusion tensor spectroscopy and imaging. Biophys J 66(1):259–267. doi:10.1016/s0006-3495(94)80775-1
2. Qiu D, Tan LH, Zhou K, Khong PL (2008) Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. NeuroImage 41(2):223–232. doi:10.1016/j.neuroimage.2008.02.023
3. Wozniak JR, Lim KO (2006) Advances in white matter imaging: a review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging. Neurosci Biobehav Rev 30(6):762–774. doi:10.1016/j.neubiorev.2006.06.003
4. Ciccarelli O, Behrens TE, Altmann DR, Orrell RW, Howard RS, Johansen-Berg H, Miller DH, Matthews PM, Thompson AJ (2006) Probabilistic diffusion tractography: a potential tool to assess the rate of disease progression in amyotrophic lateral sclerosis. Brain: J Neurol 129(Pt 7):1859–1871. doi:10.1093/brain/awl100
5. Meoded A, Kwan JY, Peters TL, Huey ED, Danielian LE, Wiggs E, Morrissette A, Wu T, Russell JW, Bayat E, Grafman J, Floeter MK (2013) Imaging findings associated with cognitive performance in primary lateral sclerosis and amyotrophic lateral sclerosis. Dement Geriatr Cogn Dis Extra 3(1):233–250. doi:10.1159/000353456
6. Sato K, Aoki S, Iwata NK, Masutani Y, Watadani T, Nakata Y, Yoshida M, Terao Y, Abe O, Ohtomo K (2010) Diffusion tensor tract-specific analysis of the uncinate fasciculus in patients with amyotrophic lateral sclerosis. Neuroradiology 52(8):729–733
7. Blaschek A, Keeser D, Muller S, Koerte IK, Sebastian Schroder A, Muller-Felber W, Heinen F, Ertl-Wagner B (2013) Early white matter changes in childhood multiple sclerosis: a diffusion tensor imaging study. AJNR Am J Neuroradiol 34(10):2015–2020. doi:10.3174/ajnr.A3581
8. Temel S, Kekligkoglu HD, Vural G, Deniz O, Ercan K (2013) Diffusion tensor magnetic resonance imaging in patients with multiple sclerosis and its relationship with disability. Neuroradiol J 26(1):3–17
9. Boespflug EL, Storrs JM, Allendorfer JB, Lamy M, Eliassen JC, Page S (2011) Mean diffusivity as a potential diffusion tensor biomarker of motor rehabilitation after electrical stimulation incorporating task specific exercise in stroke: a pilot study. Brain Imaging Behav. doi:10.1007/s11682-011-9144-1
10. van der Aa NE, Northington FJ, Stone BS, Groenendaal F, Benders MJ, Porro G, Yoshida S, Mori S, de Vries LS, Zhang J (2013) Quantification of white matter injury following neonatal stroke with serial DTI. Pediatr Res 73(6):756–762. doi:10.1038/pr.2013.45
11. Wheeler-Kingshott CA, Cercignani M (2009) About "axial" and "radial" diffusivities. Magn Reson Med: Off J Soc Magn Reson Med / Soc Magn Reson Med 61(5):1255–1260. doi:10.1002/mrm.21965
12. Klawiter EC, Schmidt RE, Trinkaus K, Liang HF, Budde MD, Naismith RT, Song SK, Cross AH, Benzinger TL (2011) Radial diffusivity predicts demyelination in ex vivo multiple sclerosis spinal cords. NeuroImage 55(4):1454–1460. doi:10.1016/j.neuroimage.2011.01.007
13. Metwalli NS, Benatar M, Nair G, Usher S, Hu X, Carew JD (2010) Utility of axial and radial diffusivity from diffusion tensor MRI as markers of neurodegeneration in amyotrophic lateral sclerosis. Brain Res 1348:156–164. doi:10.1016/j.brainres.2010.05.067
14. Seal ML, Yucel M, Fornito A, Wood SJ, Harrison BJ, Walterfang M, Pell GS, Pantelis C (2008) Abnormal white matter microstructure in schizophrenia: a voxelwise analysis of axial and radial diffusivity. Schizophr Res 101(1–3):106–110. doi:10.1016/j.schres.2007.12.489
15. Song SK, Sun SW, Ramsbottom MJ, Chang C, Russell J, Cross AH (2002) Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. NeuroImage 17(3):1429–1436
16. Della Nave R, Ginestroni A, Diciotti S, Salvatore E, Soricelli A, Mascalchi M (2011) Axial diffusivity is increased in the degenerating

superior cerebellar peduncles of Friedreich's ataxia. Neuroradiology 53(5):367–372
17. Wang JY, Abdi H, Bakhadirov K, Diaz-Arrastia R, Devous MD Sr (2012) A comprehensive reliability assessment of quantitative diffusion tensor tractography. NeuroImage 60(2):1127–1138. doi:10.1016/j.neuroimage.2011.12.062
18. Landman BA, Farrell JA, Jones CK, Smith SA, Prince JL, 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. doi:10.1016/j.neuroimage.2007.02.056
19. Ni H, Kavcic V, Zhu T, Ekholm S, Zhong J (2006) Effects of number of diffusion gradient directions on derived diffusion tensor imaging indices in human brain. AJNR Am J Neuroradiol 27(8):1776–1781
20. Papadakis NG, Murrills CD, Hall LD, Huang CL, Adrian Carpenter T (2000) Minimal gradient encoding for robust estimation of diffusion anisotropy. Magn Reson Imaging 18(6):671–679
21. Jones DK (2004) The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo Study. Magn Reson Med : Off J Soc Magn Reson Med / Soc Magn Reson Med 51(4):807–815. doi:10.1002/mrm.20033
22. Hasan KM, Parker DL, Alexander AL (2001) Comparison of gradient encoding schemes for diffusion-tensor MRI. J Magn Reson Imaging JMRI 13(5):769–780
23. Heiervang E, Behrens TE, Mackay CE, Robson MD, Johansen-Berg H (2006) Between session reproducibility and between subject variability of diffusion MR and tractography measures. NeuroImage 33(3):867–877. doi:10.1016/j.neuroimage.2006.07.037
24. Lazar M, Alexander AL (2005) Bootstrap white matter tractography (BOOT-TRAC). NeuroImage 24(2):524–532. doi:10.1016/j.neuroimage.2004.08.050
25. Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TE (2006) Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. NeuroImage 31(4):1487–1505. doi:10.1016/j.neuroimage.2006.02.024
26. Gee J, Alexander D (2006) Diffusion-tensor image registration. In: Weickert J, Hagen H (eds) Visualization and processing of tensor fields. Mathematics and visualization. Springer, Berlin, Heidelberg, pp 327–342. doi:10.1007/3-540-31272-2_20
27. Zhang H, Yushkevich PA, Alexander DC, Gee JC (2006) Deformable registration of diffusion tensor MR images with explicit orientation optimization. Med Image Anal 10(5):764–785. doi:10.1016/j.media.2006.06.004
28. Schmithorst VJ, Holland SK, Dardzinski BJ (2008) Developmental differences in white matter architecture between boys and girls. Hum Brain Mapp 29(6):696–710. doi:10.1002/hbm.20431
29. Jones DK, Griffin LD, Alexander DC, Catani M, Horsfield MA, Howard R, Williams SC (2002) Spatial normalization and averaging of diffusion tensor MRI data sets. NeuroImage 17(2):592–617
30. Christensen GE, Rabbitt RD, Miller MI (1996) Deformable templates using large deformation kinematics. IEEE Trans Image Process: Publ IEEE Signal Proc Soc 5(10):1435–1447. doi:10.1109/83.536892
31. Wang Y, Gupta A, Liu Z, Zhang H, Escolar ML, Gilmore JH, Gouttard S, Fillard P, Maltbie E, Gerig G, Styner M (2011) DTI registration in atlas based fiber analysis of infantile Krabbe disease. NeuroImage 55(4):1577–1586. doi:10.1016/j.neuroimage.2011.01.038
32. Park HJ, Kubicki M, Shenton ME, Guimond A, McCarley RW, Maier SE, Kikinis R, Jolesz FA, Westin CF (2003) Spatial normalization of diffusion tensor MRI using multiple channels. NeuroImage 20(4):1995–2009
33. Vollmar C, O'Muircheartaigh J, Barker GJ, Symms MR, Thompson P, Kumari V, Duncan JS, Richardson MP, Koepp MJ (2010) Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0T scanners. NeuroImage 51(4):1384–1394. doi:10.1016/j.neuroimage.2010.03.046
34. Brown WJ, Trost SG, Bauman A, Mummery K, Owen N (2004) Test-retest reliability of four physical activity measures used in population surveys. J Sci Med sport / Sports Med Aust 7(2):205–215
35. Marenco S, Rawlings R, Rohde GK, Barnett AS, Honea RA, Pierpaoli C, Weinberger DR (2006) Regional distribution of measurement error in diffusion tensor imaging. Psychiatry Res 147(1):69–78. doi:10.1016/j.pscychresns.2006.01.008
36. Papinutto ND, Maule F, Jovicich J (2013) Reproducibility and biases in high field brain diffusion MRI: an evaluation of acquisition and analysis variables. Magn Reson Imaging 31(6):827–839. doi:10.1016/j.mri.2013.03.004
37. Salat DH, Tuch DS, Greve DN, van der Kouwe AJ, Hevelone ND, Zaleta AK, Rosen BR, Fischl B, Corkin S, Rosas HD, Dale AM (2005) Age-related alterations in white matter microstructure measured by diffusion tensor imaging. Neurobiol Aging 26(8):1215–1227. doi:10.1016/j.neurobiolaging.2004.09.017
38. Snook L, Paulson LA, Roy D, Phillips L, Beaulieu C (2005) Diffusion tensor imaging of neurodevelopment in children and young adults. NeuroImage 26(4):1164–1173. doi:10.1016/j.neuroimage.2005.03.016
39. Alexander AL, Lee JE, Lazar M, Boudos R, DuBray MB, Oakes TR, Miller JN, Lu J, Jeong EK, McMahon WM, Bigler ED, Lainhart JE (2007) Diffusion tensor imaging of the corpus callosum in autism. NeuroImage 34(1):61–73. doi:10.1016/j.neuroimage.2006.08.032
40. Bonekamp D, Nagae LM, Degaonkar M, Matson M, Abdalla WM, Barker PB, Mori S, Horska A (2007) Diffusion tensor imaging in children and adolescents: reproducibility, hemispheric, and age-related differences. NeuroImage 34(2):733–742. doi:10.1016/j.neuroimage.2006.09.020
41. Lebel C, Walker L, Leemans A, Phillips L, Beaulieu C (2008) Microstructural maturation of the human brain from childhood to adulthood. NeuroImage 40(3):1044–1055. doi:10.1016/j.neuroimage.2007.12.053
42. Morioka S, Morimoto M, Yamada K, Hasegawa T, Morita T, Moroto M, Isoda K, Chiyonobu T, Imamura T, Nishimura A (2013) Effects of chemotherapy on the brain in childhood: diffusion tensor imaging of subtle white matter damage. Neuroradiology 55(10):1251–1257
43. Anderson AW (2001) Theoretical analysis of the effects of noise on diffusion tensor imaging. Magn Reson Med : Off J Soc Magn Reson Med / Soc Magn Reson Med 46(6):1174–1188
44. Farrell JA, Landman BA, Jones CK, Smith SA, Prince JL, van Zijl PC, Mori S (2007) Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. J Magn Reson Imaging JMRI 26(3):756–767. doi:10.1002/jmri.21053
45. Seo Y (2013) Effects of different field strengths, gradient directions, and acquisitions on fractional anisotropy in diffusion tensor imaging: a tract-based spatial statistics study. Concepts Magn Reson Part B: Magn Reson Eng 43(1):41–48
46. Barnea-Goraly N, Menon V, Eckert M, Tamm L, Bammer R, Karchemskiy A, Dant CC, Reiss AL (2005) White matter development during childhood and adolescence: a cross-sectional diffusion tensor imaging study. Cereb Cortex 15(12):1848–1854. doi:10.1093/cercor/bhi062, New York, NY : 1991