Magnetic Resonance Imaging 59 (2019) 1–9

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

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

Original contribution

# Reproducibility of multi-shell diffusion tractography on traveling subjects: A multicenter study prospective

Qiqi Tong[a], Hongjian He[a,*], Ting Gong[a], Chen Li[a], Peipeng Liang[b,c,**], Tianyi Qian[d], Yi Sun[e], Qiuping Ding[a], Kuncheng Li[c,f], Jianhui Zhong[a,g]

[a] Center for Brain Imaging Science and Technology, Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrumental Science, Zhejiang University, Hangzhou, Zhejiang, China
[b] School of Psychology, Capital Normal University, Beijing, China
[c] Beijing Key Laboratory of Magnetic Resonance Imaging and Brain Informatics, Beijing, China
[d] MR Collaboration NE Asia, Siemens Healthcare, Beijing, China
[e] MR Collaboration NE Asia, Siemens Healthcare, Shanghai, China
[f] Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China
[g] Department of Imaging Sciences, University of Rochester, Rochester, NY, USA

ARTICLE INFO

*Keywords:*
Diffusion MRI
Multi-shell diffusion
Multi-center studies
Reproducibility
Track-density imaging
Structural connectome

ABSTRACT

Reproducibility of multicenter diffusion magnetic resonance imaging has drawn more attention recently due to rapidly increasing need for large-size brain imaging studies. Advanced multi-shell diffusion models are recommended for their potentials to provide variety of physio-pathological information. While previous studies have investigated the consistency of single-shell diffusion acquisition from various hardware and protocols, a well-controlled study with multi-shell acquisition would be necessary to understand the inherent factors of reproducibility from new complexity of such acquisition protocol. In this study, three traveling subjects were scanned at eight imaging centers equipped with the same type of scanners using the same multi-shell diffusion imaging protocol. Track density imaging and structure connectomes were investigated in local-scale distribution and in distal-scale connectivity, respectively. With evaluations of the coefficient of variation and the intra-class correlation coefficient, our results indicated: 1) similar to single-shell schemes, the intra-center reproducibility of multi-shell is higher than inter-center; 2) multi-shell schemes produce higher reproducibility and precision among centers compared to the single-shell schemes; and 3) in addition to the diffusion schemes, image quality and the presence of complex fiber structure could also associated with multicenter reproducibility.

## 1. Introduction

In recent years, numerous multicenter studies of magnetic resonance imaging (MRI) have been proposed for discovering the connection patterns of brain organizations and the mechanism of cognitive function in normal populations [1,2], developing and aging populations [3–6], and in various diseases [7,8]. In these large-scale projects, diffusion imaging has been considered as one of the primary non-invasive human connectome imaging techniques for its quantitative measures of tissue viability and structural connectivity [9,10]. As a premise, the reproducibility of images is key to evaluate data consistencies within and among centers [11,12].

Historically, diffusion models using single diffusion weighting along with multiple sampling directions, such as the diffusion tensor imaging

*Abbreviations:* ADC, apparent diffusion coefficient; CSD, constrained spherical deconvolution; CSF, cerebrospinal fluid; CV, coefficient of variation; DTI, diffusion tensor imaging; DW, diffusion weighted; FA, fractional anisotropy; FOD, fiber orientation distribution; GM, gray matter; HARDI, high angular resolution diffusion imaging; ICC, intra-class correlation coefficient; MP2RAGE, magnetization-prepared two rapid acquisition gradient echo; MRI, magnetic resonance imaging; MSMT, multi-shell multi-tissue; ROI, region-of-interest; SD, standard deviation; SMS, simultaneous multi-slice; SNR, signal-to-noise ratio; TDI, track-density imaging; TE, echo time; WM, white matter

* Correspondence to: H. He, Room 314, Zhouyiqing Building, Yuquan Campus, Zhejiang University, No. 38 Zheda Road, Hangzhou 310027, China.
** Correspondence to: P. Liang, 105 West Third Ring Road North, Haidian District, Beijing 100048, China.
*E-mail addresses:* tongqq@zju.edu.cn (Q. Tong), hhezju@zju.edu.cn (H. He), gongting@zju.edu.cn (T. Gong), 21615031@zju.edu.cn (C. Li), ppliang@cnu.edu.cn (P. Liang), tianyiresearch@163.com (T. Qian), yi-sun@siemens-healthineers.com (Y. Sun), qpding@zju.edu.cn (Q. Ding), cjr.likuncheng@vip.163.com (K. Li), jzhong@zju.edu.cn (J. Zhong).

https://doi.org/10.1016/j.mri.2019.02.011
Received 10 October 2018; Received in revised form 20 February 2019; Accepted 20 February 2019

(DTI) and the high angular resolution diffusion imaging (HARDI), have been tested for reproducibility across various platforms [13–18]. Nowadays, multi-shell multi-tissue (MSMT) model with multiple diffusion weightings is becoming popular for its advantage in addressing non-Gaussian diffusion microstructures and high precise fiber tracking with tissue-specific fiber orientation distribution (FOD) [19–22]. Although the precision of tractography is improved, the reproducibility of multi-shell tractography across multiple centers has rarely been evaluated comparing with the single-shell scheme.

Fundamentally, multicenter reproducibility is likely to suffer inaccuracy and biases either from variety of MRI hardware or from inconsistent acquisition protocols [16,17,23,24]. For example, the signal-to-noise ratio (SNR) on diffusion weighted (DW) images has been a crucial factor that would affect diffusion-related measures, either diffusion-derived metrics [25–27] or FOD [28]. Echo time (TE) could also be a interfering factor for diffusion-derived tissue parameters [29]. Furthermore, while measurements with higher b-values (*e.g.* 2000 s/mm$^2$ or even 3000 s/mm$^2$) would be more precise in resolving FODs in complex tissue, they could also be a substantial challenge due to gradient hardware limit [30,31]. Some studies have been carried out without separating these factors, and noticed larger inter-scanner than intra-scanner variability with mixing hardware and protocols [14,17,32,33]. However, as far as we were aware, there have been no controlled multicenter studies to evaluate sole effect of different diffusion weighting schemes, independent of other compounding factors. This is particularly important as the diffusion gradients are a core part of multi-shell diffusion MRI.

In order to address this issue, the current study designed a well-controlled experiment, by using scanners with same type of hardware and the same version of software in eight centers. Three traveling subjects were scanned with the same protocol and operator, and their data were acquired and processed with the same operational pipeline. By controlling all these factors, it is possible to resolve the sole contribution of acquisition scheme in multicenter reproducibility. The reproducibility of both local- and distal-scale quantitative tractography features were investigated at both voxel- and region-of-interest (ROI)-based levels through Track-density imaging (TDI) [34,35] and whole-brain networks.

## 2. Methods

### 2.1. Data acquisition

The *in vivo* MR data were acquired from eight scanners located at eight centers (noted as Centers A to H). The scanners are all 3 T MR MAGNETOM Prisma (Siemens, Erlangen, Germany), equipped with the same software version of Syngo MR VD13D and gradient system with max gradient strength of 80 mT/m and slew rate of 200 T/m/s. Three healthy volunteers (one male, 23 years of age, and two females, 26 and 23 years of age) who were experienced in MR scan, and have good performance in controlling head motion, were scanned at all centers using a 64-channel head coil within one month, under the same protocol. In addition, the experiment was repeated three times in Center H (hereto referred to as Reference Center) for an intra-center investigation (episodes noted as H-1, H-2 and H-3). The study was conducted with approval from the institutional review board of Xuanwu Hospital, Capital Medical University, Beijing, China, and all volunteers have signed the informed consent forms beforehand.

MR images were acquired using the auto-align function in a Brain Dot Engine to minimize the registration error. Anatomical T1-weighted images were acquired using a 3D magnetization-prepared two rapid acquisition gradient echo (MP2RAGE) sequence (TR/TE = 5 s/2.9 ms, TI = 700, 2500 ms, FOV = 211 × 256 × 256 mm$^3$, voxel size = 1.2 × 1 × 1 mm$^3$, bandwidth = 240 Hz/Px, iPAT factor = 3, and TA = 8 mins). DW images were obtained using a simultaneous multi-slice (SMS) diffusion EPI prototype sequence (TR/TE = 5.4 s/71 ms, FOV = 220 × 220 mm, slice number = 93, voxel size = 1.5 × 1.5 × 1.5 mm$^3$, bandwidth = 1712 Hz/Px, iPAT factor = 2, and SMS factor = 3) [36], with opposite phase encoding directions (AP and PA) separately. The diffusion scheme, containing 30 vectors with uniform angular coverage on each shell (b-value = 1000, 2000, and 3000 s/mm$^2$, non-colinear between any two shells), was generated from a multiple q-space sampling tool [37]. Six non-diffusion volumes were equally separated in the scheme for the motion estimation. The total acquisition time for DW images was 19 min.

To compare the single- and multi-shell diffusion strategies, the full diffusion data were regrouped by b-values. Each three-shell data were split into three single-shell subsets, denoted as $DW_1$, $DW_2$, and $DW_3$, respectively related to the shells with b-value of 1000, 2000, and 3000 s/mm$^2$. Each of two single-shell subsets was subsequently combined to form the three two-shell datasets, denoted as $DW_{12}$, $DW_{23}$, and $DW_{13}$. These datasets as well as the full three-shell data denoted as $DW_{full}$ were further processed individually.

### 2.2. Image processing

The T1-weighted images were uniformly denoised [38] by combining two readouts from the magnetization-prepared two rapid acquisition gradient echo (MP2RAGE) sequence. Automated refined GM segmentation for the T1-weighted images was processed using Freesurfer (Athinoula A. Martinos Center for Biomedical Imaging, Harvard-MIT, US) for intensity normalization and cortical parcellation with a Desikan-Killiany template. The DW images were preprocessed using FSL (FMRIB software library, University of Oxford, UK) for field map estimation, distortion correction, and head motion correction. Subsequently, linear transform was applied to the T1-weighted image for each subject to co-register with the diffusion images.

Fiber tractography was generated with the MRtrix3 software package (Brain Research Institute, Melbourne, Australia). For single-shell DW data, FOD was calculated using the conventional constrained spherical deconvolution (CSD) method [39]. For multi-shell datasets, response functions were individually calculated from three tissues masks of white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF), which were extracted from the T1-weighted images. MSMT-CSD was then used to estimate the multi-tissue FOD with a maximum harmonic order of six [22]. Subsequently, for all the single- and multi-shell datasets, three million streamlines were calculated by anatomically-constrained tractography at a step size of 0.15 mm, with a length limitation from 3 to 250 mm and a cut-off FOD amplitude of 0.06 [40]. Lastly, spherical-deconvolution informed filtering was used to improve the accuracy of fiber reconstruction across subjects, where the normalized track density and the FOD within each voxel along the entire streamline was compared to form a cost function to determine the removal of each streamline using gradient descent optimization [41–43]. About one million streamlines were retained from each diffusion dataset.

### 2.3. TDI and connectome generation

The tractography was discretized and mapped onto a Cartesian grid to form the track density images [35]. To precisely analysis the original data, the grid resolution was set to be same with the DW images. Subsequently, the obtained TDI images were individually registered to the ones from the H-1 scan.

Tractography-based connectome matrices were generated. The template was registered into individual T1-weighted images with 12 degrees of freedom. These templates included Yeo2011 with two nodal scales [44] and the higher resolution version Schaefer2017 with four scales [45], Desikan-Killiany [46], Brodmann [47], AAL [48], Fan2016 [49], Gordon2014 [50], and Glasser2016 [51]. All the parcellations have been separated into the left and right hemispheres. The links connecting every two regions were normalized using the following

equation to avoid linear bias [52]:

$$normC_{i,j} = \frac{2}{(V_i + V_j)} \sum_{f \in C_{i,j}} \frac{1}{l(f)} \quad (1)$$

where $i$ and $j$ denote two regions, $V$ is the volume of the region, and $l$ is the length of the fibers. These weighted undirected connectome matrices were subsequently compared for reproducibility.

### 2.4. Statistics for the reproducibility evaluation

The coefficient of variation (CV) is a commonly used estimator in multicenter comparisons. It was appropriate to estimate the variability of TDI maps from multiple scans at voxel level. However, due to the intrinsic structural differences among subjects, the TDI maps would suffer severe deformations and intensity loss during inter-subject registration, so the CV was only calculated within-subject using the following equation:

$$CV = \frac{\sigma(TDI_n)}{\overline{TDI_n}} \quad (2)$$

where $\sigma$ denotes the standard deviation and $n$ denotes the scans. The intra- and inter-center CV maps from each subject were generated. These CV maps were then compared between single- and multi-shell strategies for the reproducibility of TDI results.

The intra-class correlation coefficient (ICC) uses the repeated-measure mixed effects ANOVA model to estimate the degree of reproducibility among inter- and intra-center scans, excluding the subjects' variance [53,54]. A value higher than 0.5 would indicate that the variance among scans was lower than that among subjects. In this study, the connectome matrices were generated based on the cortical parcellations of each subject, thus the matrices were minimal in registration issues between subjects. Inter-center ICC was calculated with a two-way random effects model, and intra-center ICC was calculated with a two-way mixed effects model. The formula was the same in both cases for obtaining the absolute agreement by

$$ICC = \frac{MS_{within} - MS_{\text{error}}}{MS_{within} + (MS_{between} - MS_{\text{error}})/n} \quad (3)$$

where the *MS* is the mean square, the subscripts of "*within*" and "*between*" denote within-subject and between-subject, respectively, and $n$ denotes the number of involved scans. In the matrices, the links without connections were not included in the calculation. The ICC from the intra- and inter-center scans would demonstrate the reproducibility from one center and across centers, respectively. Notably, the ICC value higher than 0.6 was labeled as substantial or more [54], and is therefore used an estimator of high reproducibility in this work. Further, the reproducibility of connectome of single- and multi-shell strategies could be compared using the ratio of high reproducibility links.

It was notable that there were more than twice samples for inter-center than for intra-center scans, thus not appropriate to directly compare the reproducibility for the two cases. Therefore, a nonparametric bootstrap sampling was performed prior for the eight inter-center scans, by randomly selecting three scans to obtain the CV and ICC value. With 1000 repeated sampling, the CVs and ICCs were averaged to represent the final inter-center results. On the other hand, the intra-center results were directly calculated from the three intra-center scans. Inter-center and intra-center reproducibility can thus be compared in our analysis.

### 2.5. ROI-based and voxel-wised comparison

ROI-based analysis was preferred when comparing the inter-subject results such as TDI maps. ROIs were manually selected for three subjects at some representative ROIs known to have primarily single-directional fibers, such as the genu (gcc) and splenium of the corpus callosum (scc), the posterior limb of the internal capsule (plic), as well as some ROIs with complex-directional fibers, such as that comprising the forceps minor (fminor), anterior thalamic radiation (atr), and corticopontine tract (cpt), and another ROI comprising the forceps major (fmajor), posterior thalamic radiation (ptr), and inferior fronto-occipital fasciculus (ifo). About 40–50 voxels were included in each ROI, the mean track density with standard deviations (SD) of which were compared among subjects and scans.

For the intra-subject results, the registration among the scans can be precise, rendering voxel-wise analysis possible. The TDI maps were therefore compared voxel-by-voxel to estimate CVs from intra- and inter-center scans, as well as the CVs from single- and multi-shell diffusion strategies.

### 2.6. SNR simulation on the tractography

A simulation experiment was designed on a synthetic phantom to investigate how SNR on DW images would affect the final TDI at 1) single and crossing fiber bundle regions; 2) single- and multi-shell diffusion strategies. The ground truth tractography used for simulation is the FiberCup, which contains single and crossing fiber bundles [55]. It was delivered to the Fiberfox simulation toolbox to generate the DW images, given the relaxation times of WM/GM/CSF and the relevant fractions [56,57]. The acquisition parameters for simulation such as the TR/TE, voxel size, and diffusion schemes were set the same as those in the *in vivo* acquisition. Notably, random complex Gaussian noise with different variances was added in the k-space domain. SNR range was designed from 4.5 to 316 (about 13 dB to 50 dB, calculated by $SNR_{dB} = 20 \times \log_{10} SNR$) within a single bundle ROI selected in WM. This simulation was repeated ten times at each noise level to minimize random bias.

The post-processing pipeline of tractography under single- and multi-shell diffusion strategies remained same with the *in vivo* datasets except the fiber number was reduced to 50 thousand dues to the simple structure and small volume of the phantom. The changes induced by different levels of noise would demonstrate the deviation of FOD peaks and track densities. The tracking precision tendency over SNR could be compared between single and crossing fiber bundle regions. The reproducibility among single- and multi-shell diffusion strategies could also be compared.

Meanwhile, the SNRs of *in vivo* datasets were also measured for comparison. The SNR was calculated following the NEMA standard by $SNR = \sqrt{2} \times Mean(S_1, S_2)/SD(S_1 - S_2)$, where $S_{1,2}$ denoted the signals from two separate images. For each dataset, five SNRs were thus obtained from each two adjacent non-diffusion images and were averaged subsequently. Since the SNR values varied across tissues due to their different T2 values, three ROIs were manually selected, two in WM like gcc and fminor/atr/cpt, and one in pure CSF.

## 3. Results

### 3.1. Reproducibility of multicenter TDI

After well-controlled the experiment with the same type of MR hardware, and the same multi-shell protocols, the reproducibility of local fiber-tract density achieves a decent level in both intra- and inter-center cases in all three traveling subjects, with an overall CV < 15% (Fig. 1). However, the superiority of smaller CV in intra-center datasets is found in all combination of DW schemes from single-shell to multi-shell, as shown in the representative histograms of subject II (Fig. 1a and b). Similarly, for the $DW_{full}$ dataset, intra-center CVs (Fig. 1c) are visibly smaller than inter-center CVs (Fig. 1d) for all three subjects. Overall, the intra-center CV with a mean value of 10.41% is generally lower than the inter-center CV of 11.31%, and a more detailed comparison is listed in Table 1. These results suggest a slightly higher reproducibility of intra-center TDI. Using multi-tissue masks excluding pure GM and CSF, the averaged intra- and inter-center CVs at the WM-

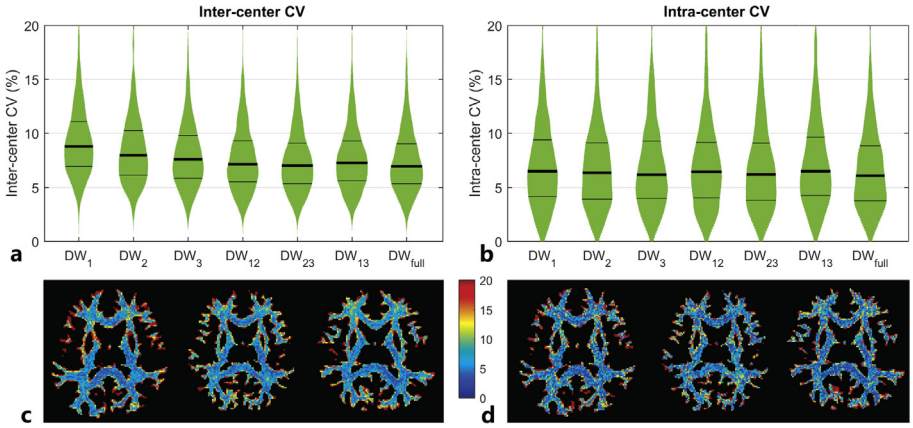


**Fig. 1.** The coefficient of variation (CV) maps and distributions of track density images. The plots in (a) and (b) are the inter- and intra-center CV distributions from different diffusion weighted datasets on subject II (the median and quantiles are marked in black lines). The subplots of (c) and (d) show the inter- and intra-center CV maps from $DW_{full}$ data on three subjects (all the maps share the same scale).

**Table 1**
The coefficient of variation (CV) distribution of track density maps in whole white matter from $DW_{full}$ dataset on three subjects.

| Subject | Inter-center CV (%) | | | Intra-center CV (%) | | |
|---|---|---|---|---|---|---|
| | Mode | Median [q1–q3][a] | Mean | Mode | Median [q1–q3] | Mean |
| I | 5.36 | 7.77 [5.25–14.90] | 12.28 | 4.45 | 7.13 [4.26–12.74] | 11.17 |
| II | 5.12 | 7.69 [5.24–12.54] | 10.63 | 4.48 | 7.02 [4.26–11.38] | 9.94 |
| III | 5.70 | 7.56 [5.38–12.72] | 10.73 | 4.62 | 6.68 [4.05–11.88] | 10.11 |

[a] Here q1 and q3 denote the first and third quartile, respectively.

GM boundary were respectively 16.68% and 19.92%, which were higher than those in pure WM regions of 6.60% and 7.96%.

The CV maps from $DW_{full}$ dataset in Fig. 1a and b shows regional differences in CV. For example, regions mainly contains single-directional fibers like gcc, scc and plic, their mean CVs were respectively 4.70% and 4.83% from intra- and inter-center scans, and complex-directional fibers, like fminor/atr/cpt, the main intra- and inter-center CVs were 6.39% and 6.60%. Moreover, further comparison of the track density across ROIs is shown in Fig. 2a. From each $DW_{full}$ data, the SD of the track density within complex-directional ROIs was 30.70 (on average), which was significantly larger than that within the single-directional ROIs (18.29). In addition, within all these ROIs, the variances of track density among the subjects were 3–4 times those in both inter- and intra-center scans. This confirmed the reproducibility of track density in a multicenter study from another aspect.

It should be noted that the tendencies of the histograms shows different between single- and multi-shell datasets in Fig. 1a and b. In the single-shell schemes, there was a trend toward a decrease in the median and quartiles of the inter-center CV when the b-value increased, whereas the intra-center trend was less obvious. In the multi-shell schemes, the CV distributions were quite similar in both the intra- and inter-center regions, but the $DW_{23}$ and $DW_{full}$ datasets showed the lowest median and quartile CVs. Similar CV distributions and tendency were observed in other two subjects.

Furthermore, the diffusion strategies also affect the track density differently among ROIs. For the H-1 scan on subject II, the mean track densities are plotted in Fig. 2b. In the single-shell schemes, the track density in all the ROIs tended to increase almost linearly with the b-value, and in the two-shell schemes, the track densities were similar in $DW_{12}$ and $DW_{23}$ but increased to the peak in the $DW_{13}$ dataset. Notably, among the four single-directional ROIs, the variance of mean track density was larger in the multi-shell schemes than in the single-shells, whereas the two complex-directional ROIs showed much smaller variance across all the datasets.

### 3.2. Reproducibility of the structural connectome

With the cortical parcellation of AAL template in the connectome matrices, the generated ICC maps in Fig. 3 showed a generally higher intra-center reproducibility of connectome matrices than the inter-

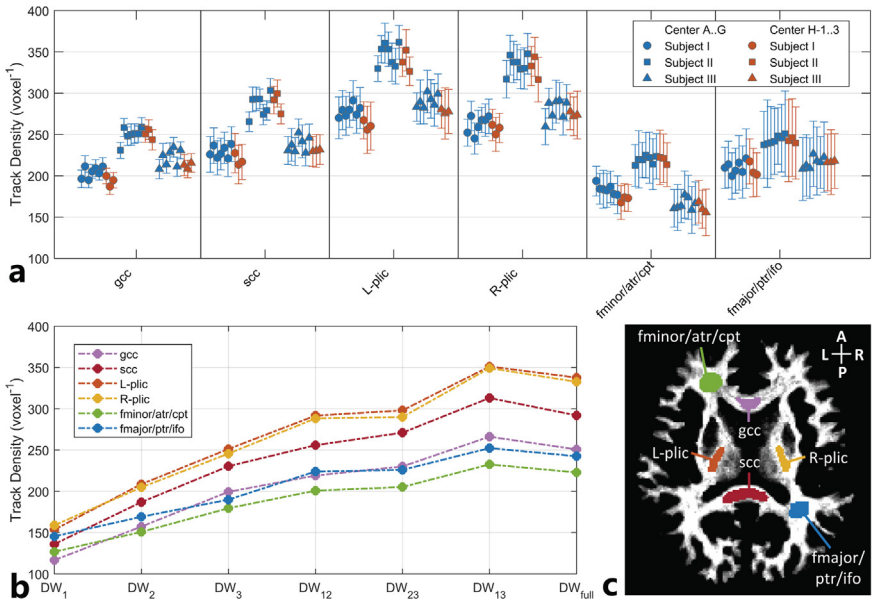


**Fig. 2.** The region-of-interest (ROI)-based track density. (a) Mean track densities generated from different diffusion weighted (DW) schemes of the H-1 scan on subject II. Different colors denote the specific ROIs defined in (b), where the ROIs are highlighted on the track density map. The genu of the corpus callosum (gcc): violet; the splenium of the corpus callosum (scc): red; the left and right branches of the posterior limb of the internal capsule (L-plic, R-plic): orange and yellow, respectively; the mixed region including the forceps minor, anterior thalamic radiation and corticopontine tract (fminor/atr/cpt): green; and another mixed region including the forceps major, posterior thalamic radiation and inferior fronto-occipital fasciculus (fmajor/ptr/ifo): blue. (c) The mean track density and related standard deviation for each ROI from $DW_{full}$ data at different centers on three subjects. The blue symbols denote the scans from Centers A to G, and the red symbols denote the three repeat scans in Center H. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

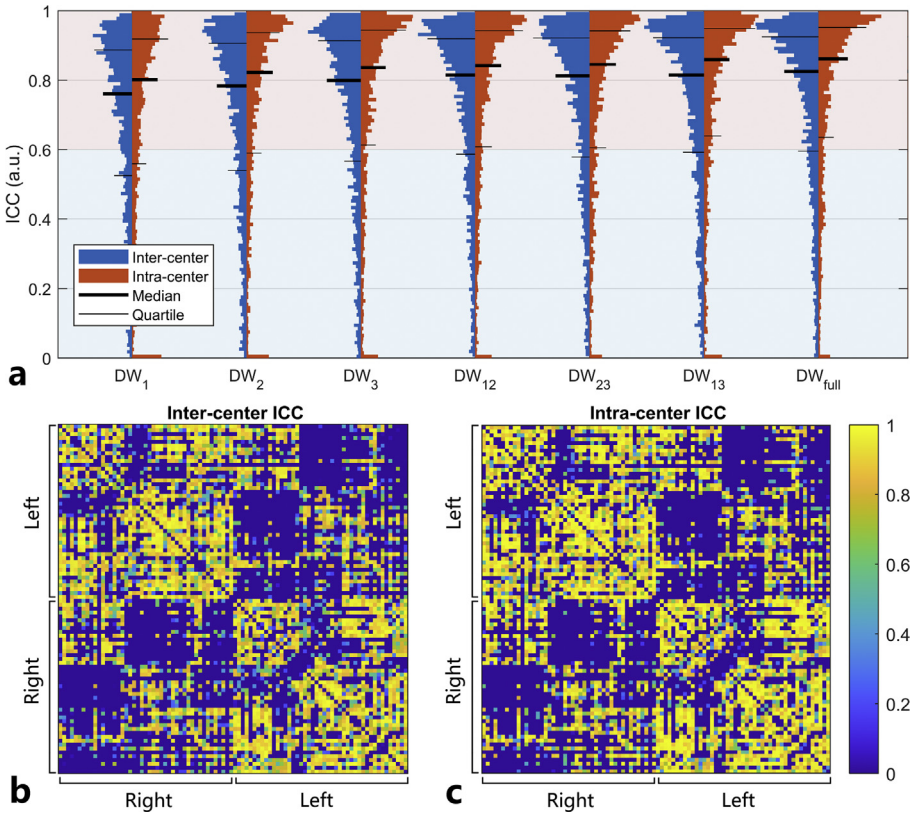


**Fig. 3.** The intra-class correlation coefficient (ICC) of the connectome matrices based on the AAL template. (a) The distributions of the inter- and intra-center ICC maps are on the left and right sides, respectively, for each DW scheme. The first quartile, median, and third quartile are marked out in black lines. (b) and (c) are ICC maps respectively generated from inter- and intra-center $DW_{full}$ data, and the nodes are classified into the "left" and "right" hemispheres.

center. From the ICC distributions of all datasets in Fig. 3a, more links in the intra-center ICC were highly reproducible than those in the inter-center. These intra- and inter-center ICC comparisons are also validated with other templates.

With different template resolutions and parcellations, the ratios of high reproducible links showed distinct differences, as shown in Table 2. Both single- and multi-shell datasets showed similar tendency. Take $DW_{full}$ dataset as an example, comparing the two Yeo2011 templates and four Schaefer2017 templates, the larger scaled template showed a decreased ratio by almost 15%. For the templates such as Fan2016, Gordon2014, and Glasser2016 with over 200 parcellations, the ratios were also reduced to about 82%. Moreover, comparing the Desikan-Killiany, Brodmann, and AAL templates at similar scales of about 80, the ratios varied from 84 to 87%, indicating that different parcellations would produce diverse reproducibility.

In the single-shell schemes, the quartile and median lines showed an increased tendency either in the inter- or intra-center distributions, demonstrating that more links from $DW_3$ dataset were highly reproducible. In the multi-shell schemes, the quartiles were slightly higher in the $DW_{13}$ and $DW_{full}$ datasets than other two datasets. Moreover, comparing the ratio of high reproducible links in Table 2, the $DW_{13}$ and $DW_{full}$ datasets also generally showed highest values in all templates. Besides, templates with more parcellations would show higher increment in single-shell datasets from $DW_1$ to $DW_3$ than those in multi-shell datasets from $DW_{12}$ to $DW_{full}$.

### 3.3. Effects of image SNR on reproducibility

With the control of different noise levels, the correlations between SNR and FOD or track density were both tested within different ROIs. In the simulated $DW_{full}$ data, the peak FOD over the SNR within multiple ROIs is shown in Fig. 4a and b. The angle difference depicted the angle

**Table 2**
Comparison of high-reproducible ratio among parcellation templates and diffusion strategies in inter-center scans.

| Template | Number of node | Ratios of high-reproducible links[a] (%) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $DW_1$ | $DW_2$ | $DW_3$ | $DW_{12}$ | $DW_{23}$ | $DW_{13}$ | $DW_{full}$ |
| Yeo2011 | 14 | 91.9 | 95.3 | 95.5 | 98.9 | 96.7 | 97.8 | 96.7 |
| | 34 | 86.7 | 87.9 | 90.1 | 91.1 | 92.4 | 91.1 | 90.4 |
| Schaefer2017 | 100 | 78.2 | 81.9 | 83.8 | 84.8 | 85.7 | 85.5 | 85.8 |
| | 200 | 72.1 | 78.8 | 79.9 | 81.4 | 82.8 | 83.4 | 83.5 |
| | 300 | 71.4 | 76.6 | 78.4 | 80.2 | 81.1 | 81.8 | 82.4 |
| | 400 | 69.2 | 75.2 | 77 | 79.5 | 80.3 | 80.9 | 81.9 |
| Desikan-Killiany | 82 | 76.4 | 81.4 | 82.9 | 81.8 | 84.2 | 84.2 | 84.1 |
| Brodmann | 82 | 79.3 | 82.1 | 83.9 | 85.1 | 86.3 | 87.1 | 86.8 |
| AAL | 90 | 78.3 | 83.4 | 83.5 | 83.6 | 84.5 | 85.6 | 86.6 |
| Fan2016 | 246 | 72.7 | 78.2 | 79.8 | 80.9 | 81.9 | 82.7 | 83 |
| Gordon2014 | 333 | 69 | 75.7 | 78.5 | 79.6 | 80.5 | 81.3 | 82.4 |
| Glasser2016 | 360 | 70.6 | 76.4 | 78.6 | 80.1 | 80 | 81.1 | 82.2 |

Abbreviation: DW = diffusion weighted.
[a] The ratio of links with ICC higher than 0.6.

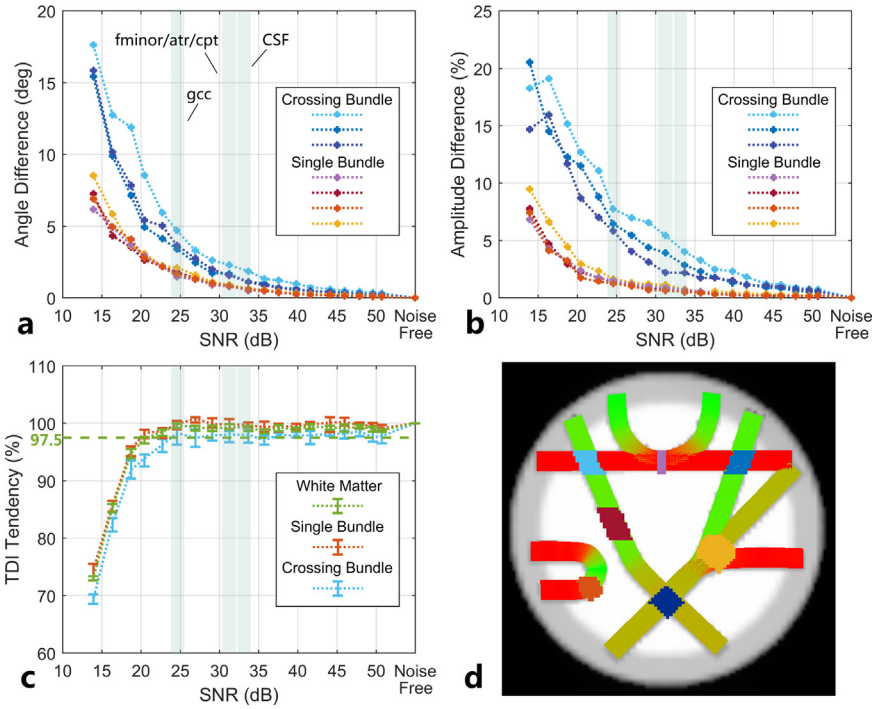


**Fig. 4.** The tendencies of the peak fiber orientation distribution (FOD) and track density over signal-to-noise ratio (SNR) within different ROIs in simulation. The angle and amplitude differences of the peak FOD are plotted in (a) and (b), respectively. The FODs in each voxel are compared to the noise-free voxels. The colors represent the specific ROIs in (d). (c) The relative tendency of the track density over the SNR. The single bundle, crossing bundle ROIs, and whole WM are plotted in blue, orange, and green, respectively. Their mean track densities with CV values over ten repetitions are plotted as error bars. A guideline of 97.5% is drawn in the green dashed line, and the SNR ranges of gcc, fminor/atr/cpt, and CSF of *in vivo* acquisition are drawn as the transparent light green rectangles. (d) The FiberCup phantom. The simulated noise-free DW image (with CSF in white, GM in gray) is overlapped with ground truth tractography, which is in direction-encoded (RGB) colors. The ROIs of the single bundle and crossing bundles are overlaid in different colors. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

between the noise-affected and the noise-free peak vectors of FOD. The amplitude difference has been normalized to the amplitude of the noise-free peak FODs. In the lower SNR cases (*i.e.* lower than 20 dB), the bias of the FOD increased nearly exponentially. The FODs changed severer in crossing bundles than single bundles, indicating that FOD with two peaks was more possible to be affected by noise than that with only one peak.

Similar tendencies were observed in the track density within single- and crossing-bundle ROIs (Fig. 4c). The mean track density flattened at high SNR around 14.1 (23 dB), but decreased rapidly below a "threshold" SNR at 10 (20 dB). In addition, the track density decreased much faster in crossing-bundle regions than in single-bundle ROIs. These results indicated a lower noise tolerance in tracking the crossing fiber than the single fiber. Therefore, in order to preserve reliable track density within the entire WM, a minimum SNR of 14.1 (23 dB) was defined for the three-shell DW scheme to ensure 97.5% track density retained with acceptable reproducibility. In practice, the SNRs in gcc, fminor/atr/cpt, and CSF for the *in vivo* acquisition were 18.04 ± 0.90 (25.07 ± 0.44 dB), 36.83 ± 3.76 (31.23 ± 0.93 dB), and 47.08 ± 4.80 (33.17 ± 0.82 dB), respectively, all of them above the simulated minimum SNR.

Comparing with the single-shell diffusion schemes, the multi-shell schemes showed higher tolerances against noise in both peak FODs and track densities in Fig. 5. When SNR was lower than the "threshold", the deviation of peak FODs' angle and amplitude increased rapidly in single-shell scheme with low b-value, especially in $DW_1$. The $DW_3$ scheme showed similar tendency with the multi-shell schemes. In the multi-shell schemes, the $DW_{12}$ scheme showed largest deviation, but was still smaller than the single-shell schemes of $DW_1$ and $DW_2$. The tendencies of track density over SNR in Fig. 5c depicted an obvious gap between single- and multi-shell schemes. In the single-shell schemes, the "threshold" SNR was higher in lower b-value cases like $DW_1$, but the track density from $DW_3$ scheme decreased the most rapidly. In the multi-shell schemes, the differences among schemes at each SNR level were smaller than the single-shell schemes.

## 4. Discussion

### 4.1. General discussion

As part of a multicenter collaboration project, this study aimed to evaluate the reproducibility of multi-shell diffusion imaging across all the involved scanners. The main objective of the present work is to evaluate the multi-shell diffusion reproducibility comparing to the single-shell which has been widely used. In addition, the study also aimed to explore possible benefits of better accuracy of multi-shell tractography against noise in tracking the single and crossing fibers, by investigating the FOD and track density from simulation.

The data acquired from the same type of scanner and a consistent protocol on traveling subjects allowed for use of an optimal post-processing pipeline to track the fibers. The reproducibility of tractography was then comprehensively estimated in both local- and distal-scale structures, respectively by the CV of TDI and ICC of the structural connectome matrix. Results from both CV and ICC demonstrated higher

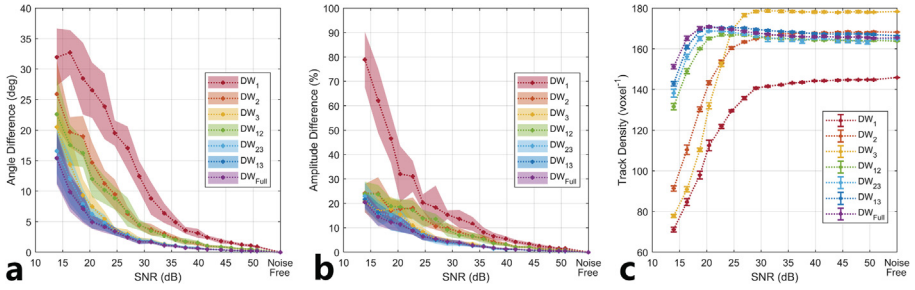


**Fig. 5.** The tendencies of the peak FOD and track density over SNR from different diffusion strategies. The ROI include all the ROIs drawn in Fig. 4d. The plots (a) and (b) present the angle and amplitude differences of the peak FOD, respectively. The mean differences with SDs over ten repetitions are shown in dots with semitransparent patches. (c) The tendency of track densities over SNR with different diffusion strategies. Mean values with SDs over ten repetitions are plotted in dots with error bars.

intra-center reproducibility than inter-center, and it was consistent across subjects. Comparing to the single-shell scheme, the multi-shell scheme showed its advantage in higher reproducibility, and more tolerance to noise in both TDI and structural connectome.

Moreover, the image quality of the SMS diffusion sequence was conducted in all centers. The type of MR scanner in this study is equipped with powerful gradient of 80 mT/m and slew rate of 200 T/m/s, with the advantages in achieving high b-value measurements at higher spatial resolution without large distortion. A homogeneous phantom ($NiSO_4$ and NaCl solution) was scanned first for an artifact test at each center. Images were acquired with the same parameters as the *in vivo* scan, and the non-diffusion images were used for investigation and measure. The percent-signal ghosting for the Nyquist ghost measure in all the centers was about 1.3–1.5%, which was lower than the ACR criteria of 2.5% [58] and was therefore acceptable for the diffusion process. Notably, neither the phantom data nor the human data were found with other visible artifacts such as blurriness, motion, or spiking in this study. Therefore, minimal artifact bias existed in the calculation for multicenter comparison.

### 4.2. Reproducibility of multi-shell TDI

Previously, the reproducibility of TDI has been tested intra- and inter-subject in a single scanner [59]. The variance of track density across subjects in our study was found to be larger than both intra- and inter-center scans in the mean track density, which is consistent with the previous study. Notably, the intra-center CV of TDI was generally lower than the inter-center CV, as shown in Fig. 1, indicating better reproducibility within one center than among different centers. The distribution of CV maps in Table 1 also confirms this difference in reproducibility, showing that more voxels had higher reproducibility in the intra-center scans, but the intra-center CV maps looks noisier due to the limited repetition scans in Center H.

Moreover, it is notable that the ROIs containing complex-directional fibers showed some differences with those containing single-directional fibers, not only in the SD of track density, but also in the CV maps. Because of the diversity in fiber directions, the fiber distribution in the complex-directional ROIs was more likely to differ from voxel to voxel, leading to a higher SD than those in the single-directional ROIs. Likewise, the CV maps also showed differences within these ROIs. This suggests that, although the acquisition was well-controlled with the same type of scanner and protocol, the reproducibility should still be re-estimated for specific tissue to study with.

Previous studies have reported that in the measurement using the diffusion tensor model, the CV of FA displayed some regional distribution characteristics, with a high CV in the low FA regions and low CV in the high FA regions [26,60]. In the current study, no such evidence was observed in the CV maps of TDI, but the CV was relatively higher at the WM-GM boundary than in the inner region. Several factors may be considered, one being the pulsation artifact. The pulsation of the CSF and vessels would deform the tissue boundary in the brain. Another factor is related to the assumption of the MSMT algorithm, where the fibers only exist within the WM; thus, at the WM-GM boundary, the number of fibers was sharply reduced. Therefore, with the existence of partial volume effect, the track density would be vastly different when the fraction of WM changes within the voxel, causing the increase in the CV.

### 4.3. Reproducibility of multi-shell structural connectome

The reproducibility of structural connectomes has been discussed *via* comparing different diffusion models [18] and tracking algorithms [61,62]. In the current study, the MSMT model was firstly used for reproducibility estimation. The generated ICC maps in Fig. 3 indicated a higher reproducibility of the structural connectome of intra-center than inter-center, which is consistent with previous multicenter studies using DTI model [63].

Moreover, the connectome matrix would also be different in terms of the scale of the cortical parcellation [64–66], and it remains to be estimated for reproducibility. Comparing different templates with various cortical parcellations resulted in large variation of the matrix values, which is consistent with previous studies. With the increase in nodes number in a template, the ICC matrix has fewer high reproducible links in the distribution. Given that the parcellation-based connectome can be affected by registration error between anatomical images and tractography at the interface of the WM-GM, a larger number of parcellations would introduce more improper segmentation of the fiber bundle [65,66]. Notably, even for parcellations with similar scales, the high reproducible ratios in Table 2 also varied. The possible reason might be that the templates were not able to cluster the fiber in the WM, generally causing a low reproducibility of some links.

### 4.4. Comparison of single- and multi-shell schemes

In general, there exists a trade-off in diffusion schemes between the accuracy of the outcome and the scan time for clinical usage. Whereas, for the HARDI scheme with single-shell, the choice of b-value is also a key factor for the precision in resolving FODs in complex tissue [30,31]. In contrast, the multi-shell scheme has the advantage in diversified coverage of q-space and resolving partial volume effects [22] comparing to the single-shell scheme with same scan time. For both TDI and structural connectome in this study, the reproducibility among all the multi-shell schemes is comparable, but the reproducibility among single-shell schemes with different b-values changes much. In the simulated data, the precision of peak FOD and track density from multi-shell schemes were generally higher than the single-shell schemes (Fig. 5). It should be noted that the diffusion direction on each shell should be balanced to minimize the variability induced by the sampling orientation [25].

For single-shell schemes, the highest track densities could be obtained from a b-value of 3000 s/mm$^2$, with lowest CV and highest ICC distribution. With the same number of diffusion directions, higher b-value results in higher amplitude of high order coefficient in the spherical harmonics, and sharper FOD in WM at the peak directions [31,67]. Therefore, high b-value would improve the precision of tractography in solving crossing fiber, and longer fibers could be reconstructed [31]. However, higher b-value scheme data were more sensitive to the noise on images. This is also demonstrated by the increase of track density in Figs. 2b and 5c.

For multi-shell schemes, similar effects were found on the reproducibility of tractography. The three-shell scheme with higher angular resolution had slightly higher reproducibility than the two-shell schemes in general. And for two-shell schemes, the $DW_{23}$ scheme showed relatively lowest CV distribution, which may indicate the highest local-scale reproducibility (Fig. 1). The $DW_{13}$ scheme showed more links with high ICC values (Fig. 3a), and may suggest a higher reproducibility in distal-scale connectivity. In the simulation, the peak FOD and track density from $DW_{13}$ scheme produced closest results to the $DW_{full}$ scheme. However, these observations require further validation, and a comprehensive study on the optimization of multi-shell scheme will be desired to achieve best accuracy and reproducibility.

### 4.5. Effects of image SNR on tractography

The simulation parameters of the DW images in this study were set strictly the same as the *in vivo* study protocol except for the artificial noise, in order to compare the image noise directly. Thus, the tendency of the FOD and track density over the SNR should reflect how accurate the tractography would be in the actual acquisition. It is obvious that when the SNR was lowest, the FOD would be largely changed, leading to the lowest track density. If the SNR increased above a 'threshold', the FOD variance would decrease and the track density would become

stable and convergent. Thus, when an acceptable rate of track density at 97.5% compared to the noise-free case was drawn, the minimum SNR in single-bundle ROI above 14.1 (23 dB) was found to be acceptable in this study. Since the SNR of *in vivo* DW images in this study were all above the minimum SNR, the accuracy of TDI from the human data should be acceptable.

Previous studies have reported that the minimum SNRs within different ROIs required for the FA calculation varied in the human brain [23,26]. In this study, although the SNR of fminor/atr/cpt ROI was higher than that of gcc in the *in vivo* data, better inter- and intra-center CVs were observed in gcc. The simulation helps to illustrate how the crossing-bundle ROIs were affected by the noise. Noise influenced the FOD amplitude and caused the crossing-bundle to be more likely to reach the stopping criteria in global tracking. Since the reconstructed tracking number was kept the same, the lengths of streamlines in low SNR were generally shorter, leading to a reduced TDI. Therefore, this characteristic reflected the fact that the regions with complex fiber bundles would be less reproducible with the existence of noise.

### 4.6. Limitations

The multicenter comparison study can be affected by some unpredictable factors from system or the subject. System characteristics such as inhomogeneity of statistic field or the gradient, gradient temperature and human effects such as head motion, cardiac pulsation, and respiration [68] are likely to be combined, and are inevitable and cannot be easily separated. In this study, cardiac gating was not included in the acquisition due to its prolonged scan time, and this might lead to more subject motion. More evidence should be sought to identify the links between these factors and the reproducibility of tractography.

The group size of three traveling subjects was relatively small to evaluate system bias among different scanners definitively. Therefore, the TDI maps were registered within subject to avoid severe variation and error induced by inaccurate inter-subject FOD registration [59]. However, a further study with large sample data would be desirable to address the issue in future.

It is worth noting that the accepted SNR range drawn in the simulated TDI results will probably be different with other protocols. The parameters used in the simulation were aimed at duplicating the actual protocols and post-processing pipeline used in this study. It was obvious that with all the parameters unaltered, the track density decreased when the SNR decreased. More tests are needed in the *in vivo* scans for the tendency, where human subjects or other diffusion phantoms can be considered, but more challenges are to be expected.

## 5. Conclusion

This work estimated the reproducibility of multi-shell diffusion images acquired from eight centers on three traveling subjects. While previous research have investigated the consistency of single-shell diffusion acquisition from various hardware and protocols, the present study was designed with a well-controlled multicenter acquisition by using the same type of scanners and the same multi-shell protocol to explore the inherent factors of reproducibility. Both track density and structural connectomes generated from fiber tractography showed higher reproducibility in intra-center than inter-center. Our further findings also suggested that a multi-shell diffusion scheme is much tolerant of noise and therefore could benefit multicenter studies with improved reproducibility than the single-shell scheme. Besides, noise in DW images would affect tractography more in solving crossing fibers than single fibers, and an acceptable SNR above 14.1 (or 23 dB) has been recommended for the multi-shell diffusion tractography.

## Acknowledgments

We would like to acknowledge participation in this work from the following hospitals and institutions in China: Beijing Neurosurgical Institute, Beijing Chaoyang Hospital, Henan Provincial People's Hospital, Meizhou People's Hospital, The First Affiliated Hospital of Zhengzhou University, Tianjin Medical University General Hospital, and the Zhongnan Hospital of Wuhan University.

*Funding*

This work was supported by the National Key R&D Program of China [grant 2017YFC0909200, 2017YFE0104000], National Natural Science Foundation of China [grant 81871428, 91632109, 81401473, 61701436, 61525106, 61473196], Beijing Talents Foundation [grant 2016000021223TD07], Beijing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding Support [grant ZYLX201609], Shanghai Key Laboratory of Psychotic Disorders (13dz2260500), and Major Scientific Project of Zhejiang Lab (no. 2018DG0ZX01).

## References

[1] Biswal BB, Mennes M, Zuo X-N, Gohel S, Kelly C, Smith SM, et al. Toward discovery science of human brain function. Proc Natl Acad Sci 2010;107:4734–9. https://doi.org/10.1073/pnas.0911855107.
[2] Liang P, Shi L, Chen N, Luo Y, Wang X, Liu K, et al. Construction of brain atlases based on a multi-center MRI dataset of 2020 Chinese adults. Sci Rep 2015;5:18216. https://doi.org/10.1038/srep18216.
[3] Jaddoe VWV, van Duijn CM, Franco OH, van der Heijden AJ, van IIzendoorn MH, de Jongste JC, et al. The generation R study: design and cohort update 2012. Eur J Epidemiol 2012;27:739–56. https://doi.org/10.1007/s10654-012-9735-1.
[4] O'Connor D, Potler NV, Kovacs M, Xu T, Ai L, Pellman J, et al. The healthy brain network serial scanning initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. Gigascience 2017;6:1–14. https://doi.org/10.1093/gigascience/giw011.
[5] Van Eijsden M, Vrijkotte TGM, Gemke RJBJ, Van der Wal MF. Cohort profile: the Amsterdam born children and their development (ABCD) study. Int J Epidemiol 2011;40:1176–86. https://doi.org/10.1093/ije/dyq128.
[6] Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TEJ, Bucholz R, et al. The human connectome project: a data acquisition perspective. Neuroimage 2012;62:2222–31. https://doi.org/10.1016/j.neuroimage.2012.02.018.
[7] Marek K, Jennings D, Lasch S, Siderowf A, Tanner C, Simuni T, et al. The Parkinson Progression Marker Initiative (PPMI). Prog Neurobiol 2011;95:629–35. https://doi.org/10.1016/j.pneurobio.2011.09.005.
[8] Toga AW, Crawford KL. The informatics core of the Alzheimer's disease neuroimaging initiative. Alzheimers Dement 2010;6:247–56. https://doi.org/10.1016/j.jalz.2010.03.001.
[9] Sporns O, Tononi G, Kötter R. The human connectome: a structural description of the human brain. PLoS Comput Biol 2005;1:0245–51. https://doi.org/10.1371/journal.pcbi.0010042.
[10] Hagmann P. From diffusion MRI to brain connectomics. 3230. 2005. p. 127. https://doi.org/10.5075/epfl-thesis-3230.
[11] Zhou X, Sakaie KE, Debbins JP, Kirsch JE, Tatsuoka C, Fox RJ, et al. Quantitative quality assurance in a multicenter HARDI clinical trial at 3 T. Magn Reson Imaging 2017;35:81–90. https://doi.org/10.1016/j.mri.2016.08.022.
[12] Marcus DS, Harms MP, Snyder AZ, Jenkinson M, Wilson JA, Glasser MF, et al. Human connectome project informatics: quality control, database services, and data visualization. Neuroimage 2013;80:202–19. https://doi.org/10.1016/j.neuroimage.2013.05.077.
[13] Zhu T, Hu R, Qiu X, Taylor M, Tso Y, Yiannoutsos C, et al. Quantification of accuracy and precision of multi-center DTI measurements: a diffusion phantom and human brain study. Neuroimage 2011;56:1398–411. https://doi.org/10.1016/j.neuroimage.2011.02.010.
[14] Vollmar C, O'Muircheartaigh J, Barker GJ, Symms MR, Thompson P, Kumari V, et al. Identical, but not the same: intra-site and inter-site reproducibility of fractional anisotropy measures on two 3.0 T scanners. Neuroimage 2010;51:1384–94. https://doi.org/10.1016/j.neuroimage.2010.03.046.
[15] Grech-Sollars M, Hales PW, Miyazaki K, Raschke F, Rodriguez D, Wilson M, et al. Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. NMR Biomed 2015;28:468–85. https://doi.org/10.1002/nbm.3269.
[16] Sasaki M, Yamada K, Watanabe Y, Matsui M, Ida M, Fujiwara S, et al. Variability in absolute apparent diffusion coefficient values across different platforms may be substantial: a multivendor, multi-institutional comparison study. Radiology 2008;249:624–30. https://doi.org/10.1148/radiol.2492071681.
[17] Teipel SJ, Reuter S, Stieltjes B, Acosta-Cabronero J, Ernemann U, Fellgiebel A, et al. Multicenter stability of diffusion tensor imaging measures: a European clinical and physical phantom study. Psychiatry Res Neuroimaging 2011;194:363–71. https://doi.org/10.1016/j.pscychresns.2011.05.012.
[18] Prčkovska V, Rodrigues P, Puigdellivol Sanchez A, Ramos M, Andorra M, Martinez-Heras E, et al. Reproducibility of the structural connectome reconstruction across diffusion methods. J Neuroimaging 2016;26:46–57. https://doi.org/10.1111/jon.

12298.
[19] Zhang H, Schneider T, Wheeler-Kingshott CA, Alexander DC. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 2012;61:1000–16. https://doi.org/10.1016/j.neuroimage.2012.03.072.
[20] Assaf Y, Basser PJ. Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain. Neuroimage 2005;27:48–58. https://doi.org/10.1016/j.neuroimage.2005.03.042.
[21] Jensen JH, Helpern JA, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: the quantification of non-Gaussian water diffusion by means of magnetic resonance imaging. Magn Reson Med 2005;53:1432–40. https://doi.org/10.1002/mrm.20508.
[22] Jeurissen B, Tournier JD, Dhollander T, Connelly A, Sijbers J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. Neuroimage 2014;103:411–26. https://doi.org/10.1016/j.neuroimage.2014.07.061.
[23] Seo Y, Wang ZJ, Morriss MC, Rollins NK. Minimum SNR and acquisition for bias-free estimation of fractional anisotropy in diffusion tensor imaging - a comparison of two analytical techniques and field strengths. Magn Reson Imaging 2012;30:1123–33. https://doi.org/10.1016/j.mri.2012.04.015.
[24] Fox RJ, Sakaie K, Lee JC, Debbins JP, Liu Y, Arnold DL, et al. A validation study of multicenter diffusion tensor imaging: reliability of fractional anisotropy and diffusivity values. Am J Neuroradiol 2012;33:695–700. https://doi.org/10.3174/ajnr.A2844.
[25] Landman BA, Farrell JAD, Jones CK, Smith SA, Prince JL, Mori S. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. Neuroimage 2007;36:1123–38. https://doi.org/10.1016/j.neuroimage.2007.02.056.
[26] Farrell JAD, Landman BA, Jones CK, Smith A, Prince JL, Van Zijl PCM, et al. Effects of SNR on the accuracy and reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 T. J Magn Reson 2010;26:756–67. https://doi.org/10.1002/jmri.21053.Effects.
[27] Jones DK, Basser PJ. "Squashing peanuts and smashing pumpkins": how noise distorts diffusion-weighted MR data. Magn Reson Med 2004;52:979–93. https://doi.org/10.1002/mrm.20283.
[28] Cohen-Adad J, Descoteaux M, Wald LL. Quality assessment of high angular resolution diffusion imaging data using bootstrap on Q-ball reconstruction. J Magn Reson Imaging 2011;33:1194–208. https://doi.org/10.1002/jmri.22535.
[29] Lin M, He H, Tong Q, Ding Q, Yan X, Feiweier T, et al. Effect of myelin water exchange on DTI-derived parameters in diffusion MRI: elucidation of TE dependence. Magn Reson Med 2018;79:1650–60. https://doi.org/10.1002/mrm.26812.
[30] Melie-García L, Canales-Rodríguez EJ, Alemán-Gómez Y, Lin CP, Iturria-Medina Y, Valdés-Hernández PA. A Bayesian framework to identify principal intravoxel diffusion profiles based on diffusion-weighted MR imaging. Neuroimage 2008;42:750–70. https://doi.org/10.1016/j.neuroimage.2008.04.242.
[31] Fan Q, Nummenmaa A, Witzel T, Zanzonico R, Keil B, Cauley S, et al. Investigating the capability to resolve complex white matter structures with high *b*-value diffusion magnetic resonance imaging on the MGH-USC Connectom scanner. Brain Connect 2014;4:718–26. https://doi.org/10.1089/brain.2014.0305.
[32] Palacios EM, Martin AJ, Boss MA, Ezekiel F, Chang YS, Yuh EL, et al. Toward precision and reproducibility of diffusion tensor imaging: a multicenter diffusion phantom and traveling volunteer study. Am J Neuroradiol 2017;38:537–45. https://doi.org/10.3174/ajnr.A5025.
[33] Pagani E, Hirsch JG, Pouwels PJW, Horsfield MA, Perego E, Gass A, et al. Intercenter differences in diffusion tensor MRI acquisition. J Magn Reson Imaging 2010;31:1458–68. https://doi.org/10.1002/jmri.22186.
[34] Calamante F, Tournier JD, Jackson GD, Connelly A. Track-density imaging (TDI): super-resolution white matter imaging using whole-brain track-density mapping. Neuroimage 2010;53:1233–43. https://doi.org/10.1016/j.neuroimage.2010.07.024.
[35] Calamante F, Tournier JD, Heidemann RM, Anwander A, Jackson GD, Connelly A. Track density imaging (TDI): validation of super resolution property. Neuroimage 2011;56:1259–66. https://doi.org/10.1016/j.neuroimage.2011.02.059.
[36] Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. Magn Reson Med 2012;67:1210–24. https://doi.org/10.1002/mrm.23097.
[37] Caruyer E, Lenglet C, Sapiro G, Deriche R. Design of multishell sampling schemes with uniform coverage in diffusion MRI. Magn Reson Med 2013;69:1534–40. https://doi.org/10.1002/mrm.24736.
[38] O'Brien K, Krueger G, Lazeyras F, Gruetter R, Roche A. A simple method to denoise MP2RAGE. Proc Int Soc Magn Reson Med 2013;21:0269. https://doi.org/10.1039/C5tc03383g.
[39] Tournier JD, Yeh CH, Calamante F, Cho KH, Connelly A, Lin CP. Resolving crossing fibres using constrained spherical deconvolution: validation using diffusion-weighted imaging phantom data. Neuroimage 2008;42:617–25. https://doi.org/10.1016/j.neuroimage.2008.05.002.
[40] Smith RE, Tournier JD, Calamante F, Connelly A. Anatomically-constrained tractography: improved diffusion MRI streamlines tractography through effective use of anatomical information. Neuroimage 2012;62:1924–38. https://doi.org/10.1016/j.neuroimage.2012.06.005.
[41] Smith RE, Tournier JD, Calamante F, Connelly A. The effects of SIFT on the reproducibility and biological accuracy of the structural connectome. Neuroimage 2015;104:253–65. https://doi.org/10.1016/j.neuroimage.2014.10.004.
[42] Smith RE, Tournier JD, Calamante F, Connelly A. SIFT: spherical-deconvolution informed filtering of tractograms. Neuroimage 2013;67:298–312. https://doi.org/10.1016/j.neuroimage.2012.11.049.
[43] Yeh CH, Smith RE, Liang X, Calamante F, Connelly A. Correction for diffusion MRI fibre tracking biases: the consequences for structural connectomic metrics. Neuroimage 2016;142:150–62. https://doi.org/10.1016/j.neuroimage.2016.05.047.
[44] Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 2011;106:1125–65. https://doi.org/10.1152/jn.00338.2011.
[45] Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 2017;28:3095–114. https://doi.org/10.1093/cercor/bhx179.
[46] Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31:968–80. https://doi.org/10.1016/j.neuroimage.2006.01.021.
[47] Brodmann K. Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Springer; 1909. https://doi.org/10.1097/00005053-191012000-00013.
[48] Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002;15:273–89. https://doi.org/10.1006/nimg.2001.0978.
[49] Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. The human Brainnetome atlas: a new brain atlas based on connectional architecture. Cereb Cortex 2016;26:3508–26. https://doi.org/10.1093/cercor/bhw157.
[50] Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. Cereb Cortex 2016;26:288–303. https://doi.org/10.1093/cercor/bhu239.
[51] Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multi-modal parcellation of human cerebral cortex. Nature 2016;536:171–8. https://doi.org/10.1038/nature18933.
[52] Hagmann P, Kurant M, Gigandet X, Thiran P, Wedeen VJ, Meuli R, et al. Mapping human whole-brain structural networks with diffusion MRI. PLoS One 2007;2:e597https://doi.org/10.1371/journal.pone.0000597.
[53] McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30–46. https://doi.org/10.1037/1082-989X.1.1.30.
[54] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159. https://doi.org/10.2307/2529310.
[55] Fillard P, Descoteaux M, Goh A, Gouttard S, Jeurissen B, Malcolm J, et al. Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom. Neuroimage 2011;56:220–34. https://doi.org/10.1016/j.neuroimage.2011.01.032.
[56] Wansapura JP, Holland SK, Dunn RS, Ball WS. NMR relaxation times in the human brain at 3.0 Tesla. J Magn Reson Imaging 1999;9:531–8. https://doi.org/10.1002/(SICI)1522-2586(199904)9:4<531::AID-JMRI4>3.0.CO;2-L.
[57] Neher PF, Laun FB, Stieltjes B, Maier-Hein KH. Fiberfox: facilitating the creation of realistic white matter software phantoms. Magn Reson Med 2014;72:1460–70. https://doi.org/10.1002/mrm.25045.
[58] American College of Radiology. Phantom Test Guidance for the ACR MRI Accreditation Program. 2005.
[59] Calamante F, Smith RE, Tournier JD, Raffelt D, Connelly A. Quantification of voxel-wise total fibre density: investigating the problems associated with track-count mapping. Neuroimage 2015;117:284–93. https://doi.org/10.1016/j.neuroimage.2015.05.070.
[60] Marenco S, Rawlings R, Rohde GK, Barnett AS, Honea RA, Pierpaoli C, et al. Regional distribution of measurement error in diffusion tensor imaging. Psychiatry Res Neuroimaging 2006;147:69–78. https://doi.org/10.1016/j.pscychresns.2006.01.008.
[61] Bastiani M, Shah NJ, Goebel R, Roebroeck A. Human cortical connectome reconstruction from diffusion weighted MRI: the effect of tractography algorithm. Neuroimage 2012;62:1732–49. https://doi.org/10.1016/j.neuroimage.2012.06.002.
[62] Buchanan CR, Pernet CR, Gorgolewski KJ, Storkey AJ, Bastin ME. Test-retest reliability of structural brain networks from diffusion MRI. Neuroimage 2014;86:231–43. https://doi.org/10.1016/j.neuroimage.2013.09.054.
[63] Bonilha L, Gleichgerrcht E, Fridriksson J, Breedlove JL, Rorden C, Nesland T, et al. Reproducibility of the structural brain connectome derived from diffusion tensor imaging. PLoS One 2015;10:1–17. https://doi.org/10.1371/journal.pone.0135247.
[64] Owen JP, Ziv E, Bukshpun P, Pojman N, Wakahiro M, Berman JI, et al. Test–retest reliability of computational network measurements derived from the structural connectome of the human brain. Brain Connect 2013;3:160–76. https://doi.org/10.1089/brain.2012.0121.
[65] Cammoun L, Gigandet X, Meskaldji D, Thiran JP, Sporns O, Do KQ, et al. Mapping the human connectome at multiple scales with diffusion spectrum MRI. J Neurosci Methods 2012;203:386–97. https://doi.org/10.1016/j.jneumeth.2011.09.031.
[66] Zalesky A, Fornito A, Harding IH, Cocchi L, Yücel M, Pantelis C, et al. Whole-brain anatomical networks: does the choice of nodes matter? Neuroimage 2010;50:970–83. https://doi.org/10.1016/j.neuroimage.2009.12.027.
[67] Tournier JD, Calamante F, Connelly A. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. NMR Biomed 2013;26:1775–86. https://doi.org/10.1002/nbm.3017.
[68] Mohammadi S, Hutton C, Nagy Z, Josephs O, Weiskopf N. Retrospective correction of physiological noise in DTI using an extended tensor model and peripheral measurements. Magn Reson Med 2013;70:358–69. https://doi.org/10.1002/mrm.24467.