

# Reliability of the corticospinal tract and arcuate fasciculus reconstructed with DTI-based tractography: implications for clinical practice

Gert Kristo · Alexander Leemans · Beatrice de Gelder ·  
Mathijs Raemaekers · Geert-Jan Rutten · Nick Ramsey

Received: 17 April 2012 / Accepted: 29 June 2012 / Published online: 7 August 2012  
© European Society of Radiology 2012

## Abstract

**Objectives** To assess the reliability of diffusion tensor imaging (DTI)-based fibre tractography (FT), which is a prerequisite for clinical applications of this technique. Here we assess the test–retest reproducibility of the architectural and microstructural features of two clinically relevant tracts reconstructed with DTI-FT.

**Methods** The corticospinal tract (CST), arcuate fasciculus (AF) and its long segment (AFL) were reconstructed in 17 healthy subjects imaged twice using a deterministic approach. Coefficients of variation (CVs) of diffusion-derived tract values were used to assess the microstructural reproducibility. Spatial correlation and fibre overlap were used to assess the architectural reproducibility.

**Results** Spatial correlation was 68 % for the CST and AF, and 69 % for the AFL. Overlap was 69 % for the CST, 61 % for the AF, and 59 % for the AFL. This was comparable to 2-mm tract shift variability. CVs of diffusion-derived tract values were at most 3.4 %.

**Conclusions** The results showed low architectural and microstructural variability for the reconstruction of the tracts. The architectural reproducibility results encourage the further investigation of the use of DTI-FT for neurosurgical planning. The high microstructural reproducibility results are promising for using DTI-FT in neurology to assess or predict functional recovery.

## Key Points

- *Magnetic resonance diffusion tensor fibre tractography is increasingly used in the neurosciences.*
- *The architectural reproducibility of fibre pathways can be up to 69 %.*
- *However the microstructural variability of fibre pathways is only 3.4 % at most.*
- *The architectural reproducibility results encourage the use of DTI-FT for neurosurgery.*
- *The microstructural reproducibility results support the use of DTI-FT in neurology.*

**Electronic supplementary material** The online version of this article (doi:10.1007/s00330-012-2589-9) contains supplementary material, which is available to authorized users.

G. Kristo · B. de Gelder  
Department of Medical Psychology and Neuropsychology,  
University of Tilburg,  
Tilburg, The Netherlands

G. Kristo · G.-J. Rutten  
Department of Neurosurgery, St. Elisabeth Hospital,  
Tilburg, The Netherlands

G. Kristo (✉) · M. Raemaekers · N. Ramsey  
Rudolf Magnus Institute of Neuroscience, and Department of  
Neurology and Neurosurgery, University Medical Center Utrecht,  
Heidelberglaan 100, mail stop G.03.124,  
3584 CX Utrecht, The Netherlands  
e-mail: g.kristo@umcutrecht.nl

A. Leemans  
Image Sciences Institute, University Medical Center Utrecht,  
Utrecht, The Netherlands

**Keywords** Diffusion tensor imaging–fiber tractography ·  
Tract specific analysis · Test–retest reliability · Neurology ·  
Neurosurgery

## Introduction

Diffusion tensor imaging (DTI) combined with fibre tractography (FT) [1, 2] is increasingly used in neurosurgery and neurology to study the corticospinal tract (CST) and arcuate fasciculus (AF) in individual patients. The CST or

pyramidal tract (the primary motor cortex part) is particularly involved in the voluntary movements of distal limbs. Knowledge of its subcortical localisation is relevant for surgical planning in patients with tumours affecting the motor pathways [3, 4]. Quantification of its microstructural properties can be relevant in the study of diseases like stroke [5, 6], or neurodegenerative disorders such as multiple sclerosis [7, 8]. The AF is generally considered to be the fourth component of the superior longitudinal fasciculus (SLF) [9]—the long association fibres running within a hemisphere [10, 11] that connect the perisylvian frontal, parietal and temporal cortex [12–14]. Within the left hemisphere it is a crucial subcortical structure for normal language function, although its specific pathways are disputed [15, 16]. For this reason precise localisation of its trajectory is relevant in surgical planning [3]. Classically, it is related to dysfunctions such as conduction aphasia [17] or primary progressive aphasia [18]. The study of its microstructural properties is relevant in cognitive degenerative disorders [19, 20].

As with any technique, the reliability (i.e. test–retest reproducibility) of DTI-FT is a prerequisite for clinical applications. In neurosurgery, gross changes to the architectural organisation can be detected by investigating the spatial configuration of fibre pathways [21]. For example, displacement, infiltration or disruption of fibre pathways by the tumour can be determined [4, 22]. In neurology, tractography-derived measures are shown to correlate with disease progress, disability and functional recovery [8]. In both cases, a key issue is that less is known about how much we can trust the DTI fibre pathways and their derived measures. Since DTI-based FT is the only commercial approach made available by the major MRI vendors, there is still a need to get a firm estimate of the reliability of this technique.

Reproducibility of architectural features addresses the spatial configuration (location and volume) of the reconstructed fibre pathways and can be measured by calculating image correlation and overlap. Reproducibility of microstructural features addresses diffusion-derived measures along the reconstructed fibre pathways and can be measured by calculating coefficients of variation (CVs). Fractional anisotropy (FA) quantifies the degree to which the diffusion magnitude depends on fibre direction, and mean diffusivity (MD) quantifies the overall amount of diffusion (independent of direction). Other measures of interest are axial or longitudinal diffusivity ( $\lambda_1$ ) and radial or transverse diffusivity ( $\lambda_\perp$ ).  $\lambda_1$  represents diffusion along the principal orientation (i.e. first eigenvalue) of the diffusion tensor (DT), whereas  $\lambda_\perp$  represents the average of the second and third eigenvalues of the DT [8, 23].

To date, few studies have measured the microstructural reproducibility of the CST and these reported low variability [24–26]. A single study reported volume agreement percentages of the CST that, depending on the number of directions of diffusion encoding, ranged from 69 to 70 % when

most of the reconstructed tracts was retained (i.e. when images were thresholded at 20 % of the tracts) [27].

Estimating the precision of tracking of the AF has been largely neglected. DTI-FT of the AF has been shown to be qualitatively reproducible between subjects, regardless of differences in data acquisition and analysis techniques [28, 29]. A previous study [30] that measured the microstructural reproducibility of the AF reported low variability. However, there are no studies that have measured the architectural reproducibility of AF.

The main goal of this study was to measure the reproducibility of the CST and AF reconstruction with DTI-FT in individual subjects. Because the specific pathways of the AF are still disputed [31], two distinct protocols were used for the reconstruction of the AF: one that defines it entirely (including SLFII tracts) (AF) [9, 32], and one that defines only its temporal lobe component or long segment (AF<sub>l</sub>) [33]. For the reconstruction of fibre pathways we used a deterministic streamline tractography approach which is standard in most commercial software and still the most widely used in clinical practice [34]. Seventeen healthy subjects participated in this study with a test–retest interval of 7 weeks. For each tract, the average FA, MD,  $\lambda_1$ ,  $\lambda_\perp$ , as well as the length and volume of tracts were calculated. The CVs of FA, MD,  $\lambda_1$  and  $\lambda_\perp$  were used to assess microstructural reproducibility, and spatial correlation and ratio of fibre volume overlap were used to assess architectural reproducibility.

## Materials and methods

### Participants

Seventeen healthy subjects (9 male and 8 female; mean age 39 years, standard deviation (SD) 12 years) participated in the experiment after giving informed consent. The research protocol and consent forms were approved by the medical ethics committee of the University Medical Center Utrecht in accordance with the Declaration of Helsinki (2008). Subjects were screened for medical, neurological and psychiatric illnesses, use of medication, and metal implants. They were strongly right-handed (mean 0.85, SD 0.15) according to the Edinburgh handedness inventory [35]. The test and retest took place with an interval of about 7 weeks (range 29–109 days, mean 51 days, SD 24 days).

### Data acquisition

All images were obtained with a 3-T Philips Achieva MRI (Philips Medical Systems, Best, the Netherlands), and the imaging protocol [36] was identical for all participants and for both imaging sessions. DTIs were acquired in transverse orientation using parallel imaging sensitivity encoding

(SENSE) ( $p$  reduction=2) with the following parameters: high angular resolution gradient set of 32 weighted directions with  $b=800$  s/mm<sup>2</sup> and one  $b=0$  s/mm<sup>2</sup> image, TR=8,481 ms, TE=60 ms, echo planar imaging (EPI) factor=59, field-of-view  $224 \times 224$  mm<sup>2</sup>, 2-mm isotropic voxel size; 60 slices with no inter-slice gap covering the whole brain. Total DTI acquisition time was 6:02 min.

#### Data processing and analysis

The DTI data were processed and analysed in ExploreDTI [37] as described below. First, subject motion and eddy-current-induced geometrical distortions were corrected, and the DTI data of the second imaging session were realigned to the DTI data of first imaging session. During image realignment, the  $b$  matrix was reoriented as in Leemans and Jones [38] to correctly preserve the orientational information. Then, the diffusion tensors and subsequently the FA, axial ( $\lambda_1$ ) and radial ( $\lambda_{\perp}$ ) diffusivity were calculated using a weighted linear least-squares estimation [39]. Finally, a standard deterministic streamline FT approach was used to reconstruct the fibre pathways of interest as described in Basser et al. [34]. The FT procedure is described in more detail in the Appendix.

#### Analysis of reliability

For each microstructural measure, the CV ( $CV=SD/mean$ ) was calculated to assess reproducibility [24, 40].

The architectural reproducibility was measured on the reconstructed tracts by means of Pearson's cross-correlation and the ratio of overlap, both taking the location of the reconstructed fibre pathways into account. Each voxel of our reconstructed tracts has a value that represents the number of fibres passing through the voxel (visitation counts). Given values of all brain voxels named  $x$  for imaging session 1 and  $y$  for session 2, the discrete correlation coefficient between them was defined with the formula

$$r = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y}$$

where cov denotes the covariance between  $x$  and  $y$  voxels, and  $\sigma$  denotes their standard deviation.  $r^2$  values can range from 0 (no correlation) to 1 (perfect correlation). The ratio of overlap is a descriptive statistic defined as the ratio of the overlapping volume of both fibre bundles present in both imaging sessions ( $V_{\text{overlap}}$ ) and the sum of the individual fibre bundle volumes [41]

$$R_{\text{overlap}}^{12} = \frac{2 \times V_{\text{overlap}}}{V_1 + V_2}$$

where  $V_1$  and  $V_2$  denote the volume of the reconstructed fibre tracts in imaging sessions 1 and 2, respectively. The

ratio of overlap can range from 0 (no overlap) to 1 (perfect overlap). Overlap values were calculated for all voxels visited at least from one fibre, and for different tract visitation thresholds (10 to 50 %, in steps of 10 %). Thresholding at more than 50 % of visiting fibres would give uninformative overlap values, because only a few voxels with high visitation counts will then be present, rather than a group of voxels representing structured tracts. Fisher's  $Z$  transformation was used on the individually estimated correlation and overlap values before group-wise comparisons.

Furthermore, the effect of visitation counts on the reproducibility of voxels within the tract was established. We expected voxels with higher visitation counts to be more reproducible and, therefore, to have a higher revisitation probability. The revisitation probability can range from 0 (the voxel is never revisited) to 1 (the voxel is always revisited), and included only voxels with visitation counts present in all subjects and sessions for both fibre tracts (70 visitation counts in our case).

## Results

#### Fibre tractography results

DTI-FT results of the CST, AF and AF $l$  reconstruction for two individuals are shown in Figs. 1, 2 and 3. Note that because the specific pathways of the AF are still disputed, we cannot exclude the reconstructed fibre tracts possibly belonging to the SLFII or other components of the SLF (Fig. 2) [9, 42].

Average FA, MD,  $\lambda_1$ ,  $\lambda_{\perp}$ , tract length and tract volume results for the CST, AF and AF $l$  are summarised in Table 1. Paired sample  $t$  tests showed no significant differences between the averaged FA, MD,  $\lambda_1$ ,  $\lambda_{\perp}$ , tract length and tract volume values of the two imaging sessions for the CST, AF and AF $l$  (lowest  $P=0.14$ ).

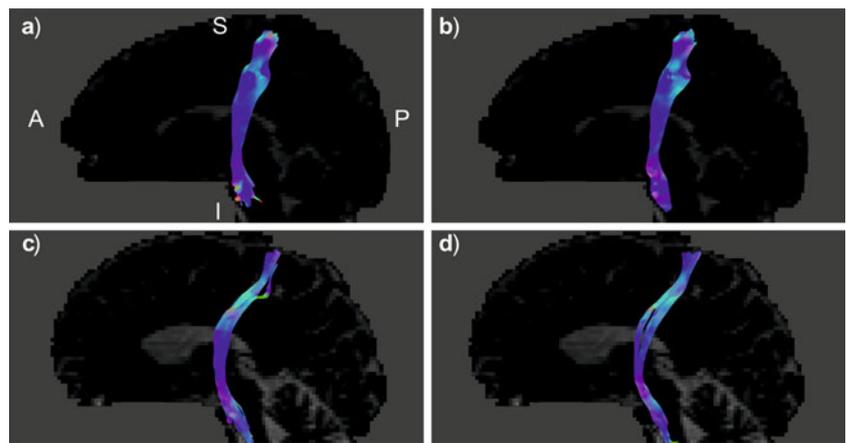
Structural differences were expected to be found between the CST and AF (and AF $l$ ) as they are functionally different from each other. Indeed, paired sample  $t$  tests showed significant differences between the averaged FA, MD,  $\lambda_1$ ,  $\lambda_{\perp}$ , tract length and tract volume values of the tracts for both imaging sessions (all  $P<0.001$ ). The CST showed higher average FA and  $\lambda_1$  values, and longer fibres, whereas the AF (and AF $l$ ) showed higher average MD and  $\lambda_{\perp}$  values, and a greater volume.

#### Reliability results

##### Microstructural reproducibility

Coefficients of variation of FA, MD,  $\lambda_1$  and  $\lambda_{\perp}$  values are summarised in Table 2. Results showed low variability in

**Fig. 1** Reconstructed corticospinal tract (CST) of subject 6 for imaging sessions 1 (a) and 2 (b), and of subject 11 for imaging sessions 1 (c) and 2 (d). The image shows the left hemisphere. *Red* is assigned to left–right, *green* to anterior–posterior (*A–P*) and *blue* to superior–inferior (*S–I*) orientations



terms of anisotropy and diffusivity for the CST, AF and AF $I$  with CVs ranging from 1.36 to 3.37 %.

Paired sample  $t$  tests showed higher  $\lambda_{\perp}$  variability ( $P=0.04$ ), and a tendency towards higher MD variability ( $P=0.06$ ) of the CST compared with the AF. Comparison of the AF $I$  with the AF showed higher  $\lambda_{\perp}$  ( $P=0.04$ ) and MD variability ( $P=0.04$ ), and a tendency towards higher  $\lambda_1$  variability ( $P=0.08$ ). No significant variability differences were found between the tracts in terms of the other microstructural measures (lowest  $P=0.12$ ).

We concluded that all tracts show high microstructural reproducibility. Microstructural reproducibility is equal for all tracts in terms of anisotropy, and different in terms of diffusivity with the AF showing the lowest variability among the tracts.

#### Tract volume reproducibility

Average CV of tract volume for the CST, AF and AF $I$  is 12.62 % (SD 10.32 %), 14.04 % (SD 11.28 %) and 16.11 % (SD 10.33 %), respectively (see Table 2). Paired sample  $t$  tests showed no significant differences in tract volume variability between the fibre bundles (lowest  $P=0.33$ ).

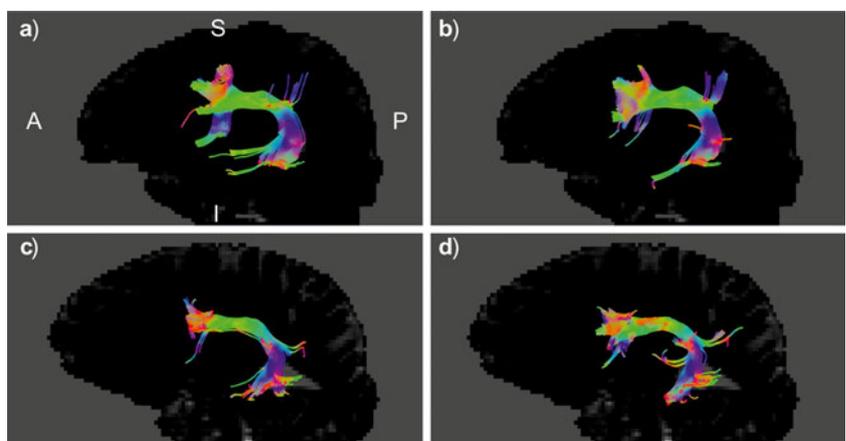
Note that the CVs of the tract volume were high for all tracts, especially compared with the CVs of the DTI measures representing the microstructural organisation of the fibre tracts.

#### Architectural reproducibility

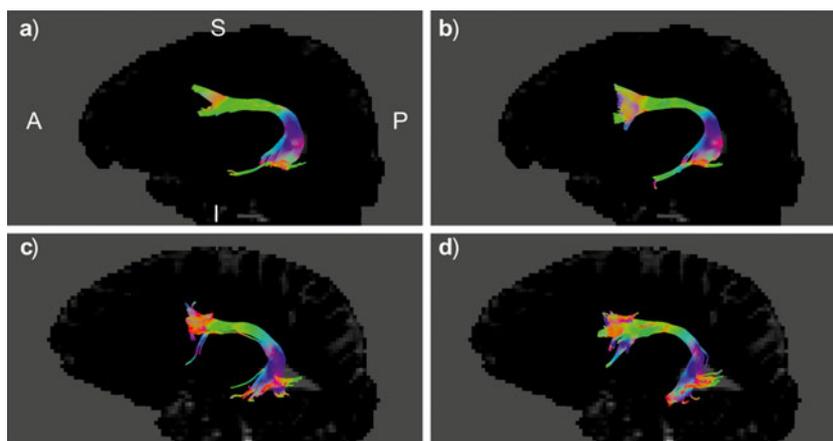
Correlation values ranged across subjects from 0.54 to 0.85 (mean 0.68, SD 0.10) for the CST, from 0.52 to 0.79 (mean 0.68, SD 0.09) for the AF, and from 0.53 to 0.80 (mean 0.69, SD 0.07) for the AF $I$ . After Fisher's  $Z$  transformation, paired sample  $t$  tests showed no significant differences between the correlation values of the tracts (lowest  $P=0.61$ ). We concluded that tracts show equal between-session correlation values.

Overlap values ranged across subjects from 0.51 to 0.81 (mean 0.69, SD 0.09) for the CST, from 0.51 to 0.68 (mean 0.61, SD 0.05) for the AF, and from 0.52 to 0.67 (mean 0.59, SD 0.05) for the AF $I$ . After Fisher's  $Z$  transformation, paired sample  $t$  tests showed higher overlap values of the CST than the AF ( $P=0.003$ ), and the AF $I$  ( $P=0.001$ ). There was no significant difference in overlap between the AF and the AF $I$  ( $P=0.19$ ). We also used different thresholds ranging from 10 to 50 % of the total number of fibres visiting voxels in both

**Fig. 2** Reconstructed arcuate fasciculus (AF) of subject 6 for imaging sessions 1 (a) and 2 (b), and of subject 11 for imaging sessions 1 (c) and 2 (d). The image shows the left hemisphere. *Red* is assigned to left–right, *green* to anterior–posterior (*A–P*) and *blue* to superior–inferior (*S–I*) orientations



**Fig. 3** Reconstructed long segment of the arcuate fasciculus (AFI) of subject 6 for imaging sessions 1 (a) and 2 (b), and of subject 11 for imaging sessions 1 (c) and 2 (d). The image shows the left hemisphere. Red is assigned to left–right, green to anterior–posterior (A–P) and blue to superior–inferior (S–I) orientations



imaging sessions. Overlap decreases at higher thresholds for all tracts (Fig. A in Electronic Supplementary Material).

To have a tangible idea of the correlation and overlap values found we made copies of the CST, AF and AFI of each imaging session and shifted them one and two voxels in x, y and z directions. Correlation and overlap values coincided in this case as they were estimated on identical copies of the tracts. Averaged correlation and overlap values (across directions, sessions and subjects) were 0.66 (SD 0.05) for one-shifted voxels, and 0.40 (SD 0.08) for two-shifted voxels for the CST. For the AF they were 0.65 (SD 0.03) for one-shifted voxels, and 0.42 (SD 0.04) for two-shifted voxels, and for the AFI they were 0.62 (SD 0.04) for one-shifted voxels, and 0.37 (SD 0.05) for two-shifted voxels. Our overlap and correlation values are similar to the ones estimated when tracts were shifted one voxel (2 mm). These results suggest a definition of tractography borders comparable to ~2 mm of uncertainty.

Lastly, Fig. 4 shows that the revisitation probability in a voxel increased with the increase in fibre visitation counts. The more fibres visited a voxel in the first imaging session, the higher the chance that the same voxel was visited in the second imaging session. Among several functions investigated, the logarithmic function fitted our data best for the CST ( $R^2=0.982$ ,  $t(1,69)=61.45$ ,  $P<0.001$ ), AF ( $R^2=0.987$ ,  $t(1,69)=71.50$ ,  $P<0.001$ ) and AFI ( $R^2=0.991$ ,  $t(1,69)=86.33$ ,  $P<0.001$ ). Remarkably, a voxel visited by only four fibres had a probability above chance (>50 %) of being revisited when fibres were again reconstructed.

## Discussion

The present study investigated the architectural and microstructural reproducibility of two clinically relevant tracts reconstructed with DTI-FT: the CST and the AF. We report

**Table 1** Values of key DTI-FT metrics in imaging sessions 1 and 2

		CST		AF		AFI	
		Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
FA	Mean	0.53	0.54	0.47	0.46	0.49	0.48
	SD	0.03	0.03	0.02	0.02	0.02	0.02
MD (mm <sup>2</sup> /s)	Mean	$7.61 \times 10^{-4}$	$7.62 \times 10^{-4}$	$7.85 \times 10^{-4}$	$7.80 \times 10^{-4}$	$7.76 \times 10^{-4}$	$7.72 \times 10^{-4}$
	SD	$2.57 \times 10^{-5}$	$1.63 \times 10^{-5}$	$2.21 \times 10^{-5}$	$2.34 \times 10^{-5}$	$2.02 \times 10^{-5}$	$2.23 \times 10^{-5}$
$\lambda_1$ (mm <sup>2</sup> /s)	Mean	$1.27 \times 10^{-3}$	$1.27 \times 10^{-3}$	$1.21 \times 10^{-3}$	$1.20 \times 10^{-3}$	$1.22 \times 10^{-3}$	$1.21 \times 10^{-3}$
	SD	$3.47 \times 10^{-5}$	$3.42 \times 10^{-5}$	$3.01 \times 10^{-5}$	$3.72 \times 10^{-5}$	$2.67 \times 10^{-5}$	$3.91 \times 10^{-5}$
$\lambda_{\perp}$ (mm <sup>2</sup> /s)	Mean	$5.08 \times 10^{-4}$	$5.06 \times 10^{-4}$	$5.72 \times 10^{-4}$	$5.68 \times 10^{-4}$	$5.52 \times 10^{-4}$	$5.51 \times 10^{-4}$
	SD	$3.07 \times 10^{-5}$	$2.72 \times 10^{-5}$	$2.39 \times 10^{-5}$	$2.09 \times 10^{-5}$	$2.18 \times 10^{-5}$	$2.14 \times 10^{-5}$
Fibre length (mm)	Mean	109.14	106.55	80.00	80.40	88.97	89.36
	SD	10.46	8.78	6.14	5.78	8.72	8.78
Fibre volume (mm <sup>3</sup> )	Mean	4,807.76	4,667.97	9,859.81	9,467.09	5,553.21	5,137.43
	SD	2,699.84	2,186.32	3,818.29	2,642.81	2,451.07	1,803.96

FA fractional anisotropy, MD mean diffusivity,  $\lambda_1$  longitudinal diffusivity,  $\lambda_{\perp}$  transverse diffusivity, CST corticospinal tract, AF arcuate fasciculus, AFI AF long segment

**Table 2** Coefficients of variation of key DTI-FT metrics

		Coefficients of variation (%)				
		FA	MD	$\lambda_{\parallel}$	$\lambda_{\perp}$	Fibre volume
CST	Mean	2.08	2.32	1.94	3.37	12.62
	SD	1.31	1.60	1.15	2.43	10.32
AF	Mean	1.94	1.36	1.37	1.89	14.04
	SD	1.29	0.87	1.00	0.96	11.28
AFI	Mean	2.41	1.81	1.71	2.60	16.11
	SD	1.46	1.30	1.33	1.46	10.33

FA fractional anisotropy, MD mean diffusivity,  $\lambda_{\parallel}$  longitudinal diffusivity,  $\lambda_{\perp}$  transverse diffusivity, CST corticospinal tract, AF arcuate fasciculus, AFI AF long segment

low architectural and microstructural variability for both tracts. Image correlation was 68 % for the CST, 68 % for the AF, and 69 % for its long segment. The overlap was 69 % for the CST, 61 % for the AF, and 59 % for its long segment. This was comparable to 2-mm tract shift variability. CVs of FA, MD,  $\lambda_{\parallel}$  and  $\lambda_{\perp}$  values were at maximum 3.4 % for the tracts.

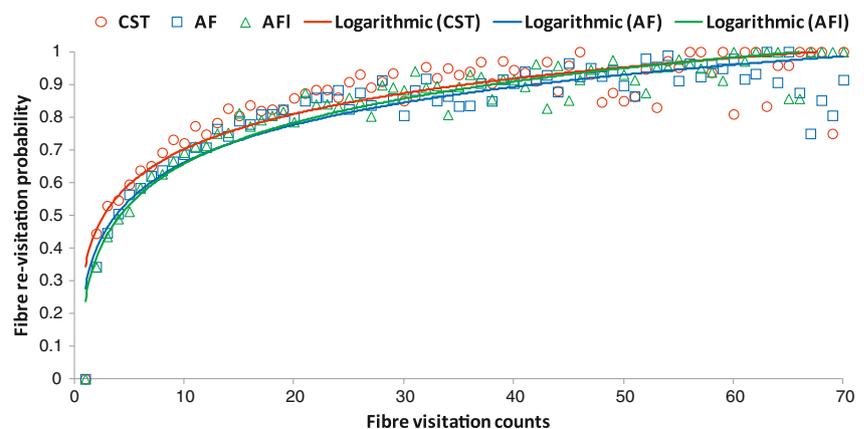
The architectural reproducibility of the fibre tracts considered did not exceed 69 %. This contrasts with a previous study that investigated the corticospinal and several other tracts and found an average overlap across all tracts between sessions of 81 % [25]. The overlap results we found may be caused by factors known to negatively affect the test–retest reliability such as changes in the position of the subject in the magnetic field of the MRI and in the radiofrequency head coil, and motion artefacts [24, 30]. Biases in the estimated diffusion measures are reduced in this study by preserving the orientational information when correcting for motion artefacts [38]. Tract volume variations may also be caused by the partial volume effect, i.e. the intra-voxel heterogeneity of different tissue types or adjacent yet differently orientated white matter fibre bundles. It emerges from the image acquisition process and is known to affect the

accuracy of DTI-FT and its estimated metrics [43]. It can directly affect the number of voxels classified as belonging to a specific white matter tract compared with another tract or grey matter [24, 44].

The architectural reproducibility results are encouraging for the further investigation of the use of DTI-FT for neurosurgical planning. The correlation and overlap values found were comparable to 2-mm tract shift variability. Although we do not assume that the tractography variability found in a test–retest investigation is a matter of tract shifting, it gives us a tangible value with which to compare our results. Our results suggest a definition of tractography borders with about 2 mm of uncertainty, which could be considered acceptable for neurosurgical planning. Of particular importance for the definition of tractography borders are results that showed voxels visited by a few fibres have a high probability of being revisited in subsequent fibre reconstruction (Fig. 4). Although these voxels may be less reliable than voxels visited by many fibres [27], our results suggest that even voxels visited by a few fibres should not be neglected in neurosurgical decision making [4]. Future studies should investigate the effect of brain lesions such as tumour or oedema on the accuracy and reliability of DTI-FT and its estimated metrics. The presence of a tumour or oedema near the tracts of interest makes the precise definition of tractography borders even more difficult. This extra variability due to the presence of a lesion is a larger problem in the paediatric population, particularly in the case of brain malformations. Whether DTI-FT is a reliable technique in certain patient populations still remains an open question.

The CST and AF show high microstructural reproducibility. The observed between-session CVs in this study are similar to those of previous studies that investigated the FA, MD [25, 30],  $\lambda_{\parallel}$  and  $\lambda_{\perp}$  [26] values of the CST and AF. These results are promising for the use of DTI-FT in clinical neurology to assess or predict functional recovery of motor and language functions [8], as these longitudinal studies are interested in changes in microstructural properties of fibre tracts

**Fig. 4** Probability values of a voxel’s fibre revisitation for the corticospinal tract (CST) (red open circles), the arcuate fasciculus (AF) (blue open squares) and its long segment (AFI) (green open triangles) for different fibre visitation counts. Data were fitted with a logarithmic function for the CST (red line), AF (blue line) and AFI (green line)



rather than in changes in their spatial configuration. However, while DTI metrics are stable when imaging is performed on the same MR system, imaging using different MR systems and various upgrades can significantly affect longitudinal DTI studies [45, 46].

The fibre tracts reconstructed here are spatially consistent with previous DTI-FT studies that investigated the CST [24–26] and the AF [28, 30, 31] (see Figs. 1, 2 and 3). The two tracts differed from each other in terms of anisotropy and diffusivity. The CST has higher average FA and  $\lambda_1$  values, whereas the AF (and AF/) has higher average MD and  $\lambda_{\perp}$  values. Although the two tracts are structurally different from each other, the microstructural reliability in terms of anisotropy and the architectural reliability in terms of image correlation were equal for the tracts. These results may reflect properties of the magnetic field. We mitigate field heterogeneity effects in this study by using the SENSE technique which has been shown to enhance DTI in the human brain at 3 T [47]. The differences found in microstructural reliability in terms of diffusivity and in architectural reliability with respect to image overlap may reflect differences in the organisation of the subcortical pathways investigated [24, 25, 30].

The generalisability of our results could theoretically be dependent on the number of gradient sampling directions. Previous research has shown that increasing the number of the directions increases the sensitivity of detecting, particularly, the branching of fibre tracts that approach the cortex [25]. However, this does not come with the benefit of increasing the test–retest reliability of these fibres. Our results are similar to those of previous studies that used 30 [30] or 32 directions [26, 27], and show similar or lower variability results than previous studies that used 60 directions [24, 25].

Our reliability results may also be dependent on the tensor model and method of fibre tracking used. The goal of this study is to measure the reliability of DTI-FT for clinical applications of this technique. For this reason we used DTI combined with a deterministic fibre tracking algorithm, which is standard in all commercial software used in clinical settings. The DTI model is well known because of its simplicity, while the deterministic fibre tracking algorithm is computationally less intensive than, for example, probabilistic methods [2]. The DTI-FT metrics found in this study are similar to those of previous studies that used a deterministic [26], continuous [30], fast marching [24] or a probabilistic tracking method [25]. Future studies should investigate whether promising fibre tracking techniques [48, 49] are also reliable.

To summarise, this study measures the architectural and microstructural reproducibility of the CST and the AF reconstruction with DTI-FT. The results showed low architectural and microstructural variability for the reconstruction of

the tracts. The architectural reproducibility results are encouraging for the further investigation of the use of DTI-FT for neurosurgical planning. The high microstructural reproducibility results are promising for using DTI-FT to study pathological processes in neurological disease [8].

## Appendix

### Fibre tractography procedure

First, whole brain tractography was performed for all individuals and imaging sessions. FA thresholds were set to 0.20 to initiate and continue tracking and the angle threshold was set to 30° for both tracts. Only fibres with a minimum length of 50 mm were considered.

Next, the CST, the AF and AF/ were extracted for each individual from regions of interest (ROIs) according to a priori information of tract location [50, 51]. The tracts of interest were only delineated in the left hemisphere considering the left hemisphere lateralisation of these tracts when subjects are strongly right-handed [52, 53]. The ROI sets of all tracts of interest were manually placed only by the first author (G.K.) as a previous study reported a high interobserver reliability in tract-specific analysis [54]. This was done following the guidelines and protocols described in previous studies, where figures showing the ROI placement are also present [9, 32, 33, 51]. The CST was extracted by placing two ROIs on axial slices. The first ROI included the cerebral peduncle at the level of the decussation of the superior cerebellar peduncle. The second ROI was drawn right after the bifurcation to the motor and sensory cortex to include only primary motor cortex, and not sensory tracts [51]. The AF was extracted by placing two ROIs on coronal slices. The first ROI was selected where the AF, appearing as a green triangular shape on coronal images (indicating anterior/posterior orientation), was seen to be largest. The second ROI was selected at the level of the splenium of the corpus callosum, where the AF makes a sharp turn towards the temporal lobe [9, 32]. The AF/ was defined by using the first ROI used for the delineation of the AF, and by placing a second ROI on an axial slice through which the AF passes in the superior/inferior direction [33].

The ROI sets of all tracts of interest were per individual the same for both imaging sessions. In doing so, the variability in ROI designation introduced by the operator is eliminated and does not affect the reliability results. However, we did not underestimate reliability because the protocol followed for tract delineation already reduced the variability in ROI designation to a minimum. All ROIs were placed subcortically [51] and were large enough to encompass all possible fibres belonging to the tracts of interest [33]. The subcortical placement of ROIs ensures proper

delineation of tracts, such as the CST, that are known to substantially differ in volume among individuals. The large size of the ROIs ensures similar tractography results for different operators.

For each tract (corticospinal and AF) per subject and imaging session, the following parameters were calculated: average FA, average MD, average  $\lambda_1$ , average  $\lambda_{\perp}$ , average length of tracts, and volume of tracts [24–26]. FA, MD,  $\lambda_1$  and  $\lambda_{\perp}$  values were obtained by averaging across all voxels in the tract and each voxel was counted only once.

## References

- Mori S, van Zijl PC (2002) Fiber tracking: principles and strategies - a technical review. *NMR Biomed* 15:468–480
- Tournier JD, Mori S, Leemans A (2011) Diffusion tensor imaging and beyond. *Magn Reson Med* 65:1532–1556
- Bello L, Gambini A, Castellano A et al (2008) Motor and language DTI Fiber Tracking combined with intraoperative subcortical mapping for surgical removal of gliomas. *NeuroImage* 39:369–382
- Kinoshita M, Yamada K, Hashimoto N et al (2005) Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation. *NeuroImage* 25:424–429
- Kunimatsu A, Aoki S, Masutani Y, Abe O, Mori H, Ohtomo K (2003) Three-dimensional white matter tractography by diffusion tensor imaging in ischaemic stroke involving the corticospinal tract. *Neuroradiology* 45:532–535
- van der Aa NE, Leemans A, Northington FJ et al (2011) Does diffusion tensor imaging-based tractography at 3 months of age contribute to the prediction of motor outcome after perinatal arterial ischemic stroke? *Stroke* 42:3410–3414
- van Hecke W, Nagels G, Leemans A, Vandervliet E, Sijbers J, Parizel PM (2010) Correlation of cognitive dysfunction and diffusion tensor MRI measures in patients with mild and moderate multiple sclerosis. *J Magn Reson Imaging* 31:1492–1498
- Ciccarelli O, Catani M, Johansen-Berg H, Clark C, Thompson A (2008) Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. *Lancet Neurol* 7:715–727
- Makris N, Kennedy DN, McInerney S et al (2005) Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb Cortex* 15:854–869
- Gloor P (1997) The temporal lobe and the limbic system. Oxford University Press, New York
- Kiernan JA (1998) Barr's the human nervous system: an anatomical view point. Lippincott–Raven, Philadelphia
- Echtalew C (1992) Correlative anatomy of the nervous system. Macmillan, New York
- Dejerine J (1985) Anatomie des Centres Nerveux. Rueff et Cie, Paris
- Nieuwenhuys RVJHC (1988) The human central nervous system. Springer-Verlag, Berlin
- Petrides M, Pandya DN (1988) Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J Comp Neurol* 273:52–66
- Bernal B, Ardila A (2009) The role of the arcuate fasciculus in conduction aphasia. *Brain* 132:2309–2316
- Benson DF, Sheremata WA, Bouchard R, Segarra JM, Price D, Geschwind N (1973) Conduction aphasia. A clinicopathological study. *Arch Neurol* 28:339–346
- Mesulam MM (2001) Primary progressive aphasia. *Ann Neurol* 49:425–432
- Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. *Cortex* 44:953–961
- Verhoeven JS, Rommel N, Prodi E et al (2011) Is there a common neuroanatomical substrate of language deficit between autism spectrum disorder and specific language impairment? *Cereb Cortex*. doi:10.1093/cercor/bhr292
- Clark CA, Byrnes T (2010) DTI and tractography in neurosurgical planning. In: Jones DK (ed) Diffusion MRI: theory, methods and applications. Oxford University Press, New York
- Nimsky C, Ganslandt O, Hastreiter P et al (2005) Intraoperative diffusion-tensor MR imaging: shifting of white matter tracts during neurosurgical procedures – initial experience. *Radiology* 234:218–225
- Wheeler-Kingshott CA, Cercignani M (2009) About “axial” and “radial” diffusivities. *Magn Reson Med* 61:1255–1260
- Ciccarelli O, Parker GJ, Toosy AT et al (2003) From diffusion tractography to quantitative white matter tract measures: a reproducibility study. *NeuroImage* 18:348–359
- 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* 33:867–877
- Danielian LE, Iwata NK, Thomasson DM, Floeter MK (2010) Reliability of fiber tracking measurements in diffusion tensor imaging for longitudinal study. *NeuroImage* 49:1572–1580
- Tensaouti F, Lahlou I, Clarisse P, Lotterie JA, Berry I (2011) Quantitative and reproducibility study of four tractography algorithms used in clinical routine. *J Magn Reson Imaging* 34:165–172
- Catani M, Howard RJ, Pajevic S, Jones DK (2002) Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage* 17:77–94
- Mori S, Kaufmann WE, Davatzikos C et al (2002) Imaging cortical association tracts in the human brain using diffusion-tensor-based axonal tracking. *Magn Reson Med* 47:215–223
- Wang JY, Abdi H, Bakhadirov K, az-Arrastia R, Devous MD Sr (2012) A comprehensive reliability assessment of quantitative diffusion tensor tractography. *NeuroImage* 60:1127–1138
- Catani M, Jones DK, Ffytche DH (2005) Perisylvian language networks of the human brain. *Ann Neurol* 57:8–16
- Matsumoto R, Okada T, Mikuni N et al (2008) Hemispheric asymmetry of the arcuate fasciculus: a preliminary diffusion tensor tractography study in patients with unilateral language dominance defined by Wada test. *J Neurol* 255:1703–1711
- Lebel C, Beaulieu C (2009) Lateralization of the arcuate fasciculus from childhood to adulthood and its relation to cognitive abilities in children. *Hum Brain Mapp* 30:3563–3573
- Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A (2000) In vivo fiber tractography using DT-MRI data. *Magn Reson Med* 44:625–632
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113
- Jones DK, Leemans A (2011) Diffusion tensor imaging. *Methods Mol Biol* 711:127–144
- Leemans A, Jeurissen B, Sijbers J et al. (2009) Explore DTI: a graphical toolbox for processing, analyzing, visualizing diffusion MR data. In: 17th Annual Meeting Proc Intl Soc Mag Reson Med, Hawaii USA, pp 3537
- Leemans A, Jones DK (2009) The B-matrix must be rotated when correcting for subject motion in DTI data. *Magn Reson Med* 61:1336–1349
- Chang LC, Jones DK, Pierpaoli C (2005) RESTORE: robust estimation of tensors by outlier rejection. *Magn Reson Med* 53:1088–1095

40. Bland JM, Altman DG (1996) Measurement error proportional to the mean. *BMJ* 313:106
41. Raemaekers M, Vink M, Zandbelt B, van Wezel RJ, Kahn RS, Ramsey NF (2007) Test-retest reliability of fMRI activation during prosaccades and antisaccades. *NeuroImage* 36:532–542
42. Bernal B, Altman N (2010) The connectivity of the superior longitudinal fasciculus: a tractography DTI study. *Magn Reson Imaging* 28:217–225
43. Alexander AL, Hasan KM, Lazar M, Tsuruda JS, Parker DL (2001) Analysis of partial volume effects in diffusion-tensor MRI. *Magn Reson Med* 45:770–780
44. Vos SB, Jones DK, Viergever MA, Leemans A (2011) Partial volume effect as a hidden covariate in DTI analyses. *NeuroImage* 55:1566–1576
45. Takao H, Hayashi N, Kabasawa H, Ohtomo K (2012) Effect of scanner in longitudinal diffusion tensor imaging studies. *Hum Brain Mapp* 33:466–477
46. Takao H, Hayashi N, Ohtomo K (2011) Effect of scanner in asymmetry studies using diffusion tensor imaging. *NeuroImage* 54:1053–1062
47. Jaermann T, Crelier G, Pruessmann KP et al (2004) SENSE-DTI at 3 T. *Magn Reson Med* 51:230–236
48. Jeurissen B, Leemans A, Jones DK, Tournier JD, Sijbers J (2011) Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution. *Hum Brain Mapp* 32:461–479
49. Hosey T, Williams G, Ansorge R (2005) Inference of multiple fiber orientations in high angular resolution diffusion imaging. *Magn Reson Med* 54:1480–1489
50. Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PC, Mori S (2004) Fiber tract-based atlas of human white matter anatomy. *Radiology* 230:77–87
51. Wakana S, Caprihan A, Panzenboeck MM et al (2007) Reproducibility of quantitative tractography methods applied to cerebral white matter. *NeuroImage* 36:630–644
52. Catani M, Allin MP, Husain M et al (2007) Symmetries in human brain language pathways correlate with verbal recall. *Proc Natl Acad Sci U S A* 104:17163–17168
53. Vernooij MW, Smits M, Wielopolski PA, Houston GC, Krestin GP, van der LA (2007) Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: a combined fMRI and DTI study. *NeuroImage* 35:1064–1076
54. Yasmin H, Aoki S, Abe O et al (2009) Tract-specific analysis of white matter pathways in healthy subjects: a pilot study using diffusion tensor MRI. *Neuroradiology* 51:831–840