

Between session reproducibility and between subject variability of diffusion MR and tractography measures

E. Heiervang,^{a,b} T.E.J. Behrens,^a C.E. Mackay,^c M.D. Robson,^c and H. Johansen-Berg^{a,*}

^aCentre for Functional Magnetic Resonance Imaging of the Brain, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford OX3 9DU, UK

^bDepartment of Child and Adolescent Mental Health, Haukeland University Hospital, 5021 Bergen, Norway

^cUniversity of Oxford Centre for Clinical Magnetic Resonance Research, John Radcliffe Hospital, Oxford OX3 9DU, UK

Received 15 March 2006; revised 14 July 2006; accepted 28 July 2006

Available online 26 September 2006

As diffusion tractography is increasingly used to generate quantitative measures to address clinical questions, it is important to characterise the inter-session reproducibility and inter-subject variability of these measures. Here, we assess the reproducibility and variability of diffusion tractography measures using diffusion data from 8 subjects scanned 3 times. We used probabilistic tractography to define the cingulum bundle, pyramidal tracts, optic radiations and genu of the corpus callosum in each individual data set using three different methods of seed definition. Measures of mean fractional anisotropy (FA) and mean diffusivity (MD) along the tracts were more reproducible than measures of tract volume. Further, tracts defined using a two region of interest (ROI) approach were more reproducible than those defined using manually placed seed masks alone. For mean FA taken from tracts defined using the two ROI approach, inter-session coefficients of variation (CV) were all below 5% and inter-subject CVs were below 10%; for mean MD inter-session, CVs were all below 3% and inter-subject CVs were below 8%. We use the variability measures found here to calculate the sample sizes required to detect changes in FA, MD or tract volume of a given size, either between groups of subjects or within subjects over time. Finally, we compare tractography results using 60 diffusion encoding directions to those found using a subset of 12 directions; the number of diffusion directions did not have a significant effect on reproducibility, but tracts derived using fewer directions were consistently smaller than those derived using 60 direction data. We suggest that 12 direction data are sufficient for reproducibly defining the core of large bundles but may be less sensitive to smaller pathways.

© 2006 Elsevier Inc. All rights reserved.

Introduction

Diffusion tensor imaging (DTI) has become a popular tool for assessing brain white matter integrity in vivo. It is a safe, non-invasive technique that is sensitive to the diffusion properties of

tissue water and provides information about orientation of fibre pathways (Basser et al., 1994). DTI can generate quantitative measures, such as fractional anisotropy (FA), that are thought to be markers of tissue microstructure (Beaulieu, 2002). As these measures are increasingly used to address clinical questions, it is important to fully characterise their reproducibility.

Fractional anisotropy (FA) quantifies the degree to which diffusion is direction-dependent; FA is low in large pools of free water, where diffusion is equal in all directions, but high in structured tissue, such as brain white matter, where diffusion is least hindered along the axis of a fibre bundle. Measures of FA have been used to investigate changes in white matter microstructure with disease (Bammer et al., 2000; Buchsbaum et al., 1998; Rose et al., 2000; Toosy et al., 2003) as well as normal ageing (Pfefferbaum et al., 2000) and development (McKinstry et al., 2002). FA can be quantified using histogram measures of the entire brain white matter (Bozzali et al., 2001; Cercignani et al., 2003; Molko et al., 2002; Rocca et al., 2003) or by averaging values from within manually defined white matter regions of interest (Griffin et al., 2001; O'Sullivan et al., 2001; Toosy et al., 2003). Mean diffusivity (MD) quantifies the overall amount of diffusion at a voxel, regardless of the direction of diffusion. MD has also proved to be a clinically useful measure, with increased MD typically accompanying decreased FA. Specific increases in MD have been reported in a number of clinical conditions (Chabriat et al., 1999; Charlton et al., 2006; Ciccarelli et al., 2003b; Foong et al., 2000; Toosy et al., 2003).

An alternative to defining white matter regions of interest, or calculating histogram measures across the whole brain white matter, is to use diffusion tractography (Behrens et al., 2003b; Jones et al., 1999b; Lazar and Alexander, 2005; Mori et al., 1999) to estimate the course of fibre pathways through the white matter and then derive measures such as FA, MD or connection probability, from within those pathways. This approach has previously been used both in clinical conditions (Abe et al., 2004; Aoki et al., 2005; Ciccarelli et al., 2006; Jones et al., 2005), and in basic neuroscience (Behrens et al., 2003a; Gong et al., 2005; Huang et al., 2005; Parker et al., 2005; Toosy et al., 2004).

* Corresponding author. Fax: +44 1865 222717.

E-mail address: heidi@fmrib.ox.ac.uk (H. Johansen-Berg).

Available online on ScienceDirect (www.sciencedirect.com).

In the face of increasing use of DTI for quantitative comparisons between subjects or over time, it is essential to assess the reproducibility of the technique (Ciccarelli et al., 2003a; Ding et al., 2003; Pfefferbaum et al., 2003). Normal variability and test–retest stability are key issues for group and longitudinal study designs yet few studies have addressed these questions. Pfefferbaum et al. (2003) reported a test–retest coefficient of variation (CV) of 1.9% for FA within the corpus callosum. It is clear from other work that variability measures can differ between tracts. Ciccarelli et al. (2003a) scanned four subjects on two occasions and found CVs of FA from within tractography-defined tracts of 6.2%, 7.1% and 5% for callosum, optic radiation and pyramidal tract, respectively. Inter-subject CV from the same study was in the range of 6–9% for these tracts (Ciccarelli et al., 2003a).

Here, we assess the reproducibility of DTI measures from histograms, ROI-based analyses and tractography with 8 subjects scanned three times with the same sequence and scanner.

Methods

Data acquisition

We acquired MR data in eight healthy adult subjects (4 men, 4 women, age range 21–34 years). All subjects were right-handed, with no history of psychiatric or neurological disease. Informed written consent was obtained from all subjects in accordance with ethical approval from the Central Office for Research Ethics Committees.

Scans were obtained on three separate days within a 3 month period on a 1.5 T Siemens Sonata MR scanner with maximum gradient strength of 40 mT m^{-1} . Diffusion-weighted data were acquired using echo planar imaging ($72 \times 2 \text{ mm}$ thick axial slices, matrix 128×104 , field of view $256 \times 208 \text{ mm}$, giving a voxel size of $2 \times 2 \times 2 \text{ mm}$). The diffusion weighting was isotropically distributed along 60 directions using a b value of 1000 s mm^{-2} .

A T1-weighted anatomical image was acquired on one session for each subject using a 3D FLASH sequence (repetition time = 12 ms, echo time = 5.65 ms and flip angle = 19° , with elliptical sampling of k space, giving a voxel size of $1 \times 1 \times 1 \text{ mm}$ in 5 min and 5 s).

Image processing

Image analysis was carried out using tools from the FMRIB Software Library (FSL, www.fmrib.ox.ac.uk/fsl; Smith et al., 2004).

Brain extraction, registration and tissue type segmentation

We skull-stripped (Smith, 2002) diffusion-weighted, T1-weighted and MNI standard brain template images (Evans et al., 2003) and performed affine registration (Jenkinson and Smith, 2001) to derive transformation matrices among the three spaces. We performed probabilistic tissue type segmentation and partial volume estimation on the T1-weighted image (Zhang et al., 2001). The white matter partial volume estimate map was thresholded at 0.3 and binarised to produce individual subject white matter masks which were realigned into standard brain space using the parameters derived above with nearest neighbour interpolation.

Diffusion tensor fitting and histogram analysis

FDT (FMRIB's Diffusion Toolbox) was used to fit a diffusion tensor at each brain voxel in the diffusion data and calculate voxel-wise values for FA and MD. Individual session FA and MD maps

were realigned into standard brain space using the affine transformation matrices derived previously and masked by each subject's own white matter mask to produce standard space maps of white matter FA and MD. For each scan session, histograms of white matter FA and MD were generated in Matlab (Version 6, MathWorks, Natick, MA) using 100 bins and histogram measures of peak height, mean and standard deviation were calculated.

ROI definition: diffusion space

Single voxel regions of interest (ROI) were manually placed on FA maps (in the original space of the diffusion images) (Fig. 1) according to the following criteria (based on Ciccarelli et al. (2003a) for tracts 2, 3, 4):

1. Cingulum bundle (CB): The sagittal slice in which the CB appeared longest was selected and the CB voxel above the body of the corpus callosum with the highest FA value was identified.
2. Genu of CC (GCC): The axial slice above the one where the genu first shows fully (i.e., without any apparent partial volume effect at the midline) was selected and the voxel closest to the midline,

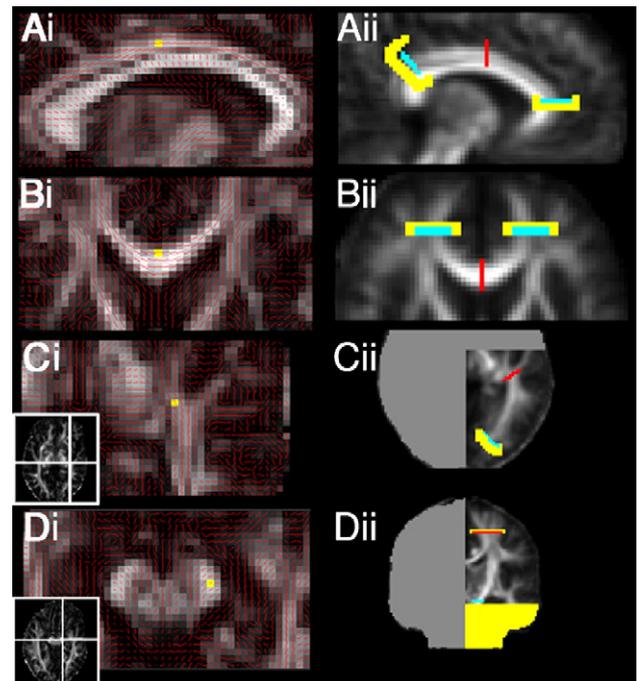


Fig. 1. Seed placement for tractography. Examples are shown for single voxel seeds in diffusion space (left column) and the two ROI technique using standard space masks (right column). Left: Single seed voxels were placed manually for each individual subject. Example images are shown for a single subject. The background image is the FA map, with estimates of principal diffusion direction at each voxel overlaid as a red vector and the single voxel ROI overlaid in yellow. ROI placement is shown for cingulum bundle (Ai, sagittal slice), genu of corpus callosum (Bi, axial slice), optic radiation (Ci, axial slice) and pyramidal tract (Di, axial slice). 3×3 voxel ROIs were created by 2D dilation (in the axial plane) of the single seed voxels shown here. For criteria used see text. (B) Examples of mask definition for the two ROI, standard space approach. The background image is the group mean FA map, overlaid with seed masks (blue), target masks (red), termination masks (yellow) and removal masks (grey). ROI placement is shown for cingulum bundle (Aii, sagittal slice), genu of corpus callosum (Bii, axial slice), optic radiation (Cii, axial slice) and pyramidal tract (Dii, coronal slice). For criteria used see text.

Table 1
Results for histogram analysis of FA and MD of whole brain white matter

Diffusion measure	Histogram measure	Mean value across subjects	Variability measures		
			Inter-session		Inter-subject
			Mean inter-session CV (%)	Inter-subject CV—based on mean across sessions (%)	Inter-subject CV—based on session one only (%)
FA	Mean	0.389	0.78	2.91	3.46
	Mode	0.384	1.55	2.99	4.92
	Peak height	0.059	0.94	4.66	4.75
	SD	0.139	0.90	4.12	3.89
MD	Mean	7.18×10^{-4}	0.99	1.66	2.04
	Mode	6.94×10^{-4}	1.48	1.91	2.64
	Peak height	0.220	3.54	6.77	6.58
	SD	0.94×10^{-4}	3.23	19.27	17.92

For each histogram measure (mean, mode, peak height, SD), we calculated the mean value across subjects (column 3) and also the inter-session and inter-subject variability in these mean values (columns 4–6).

in the most anterior row with no partial volume effect, with the highest FA value was identified.

3. Optic radiation (OR): The highest axial slice on which the cerebral peduncle, but not the anterior limb of the internal capsule, is visible was selected. On the axial slice above this slice, we overlaid the estimates of principal diffusion direction and found the voxel with the highest FA at the apex of the arc around the lateral ventricle with the main eigenvector in an anterior-medial to posterior-lateral orientation.
4. Pyramidal tract (PT): The first axial slice below the slice on which the optic tracts were visible was selected and the voxel with highest FA in the middle third of the cerebral peduncle was identified.

These individual voxels were dilated in 2D to produce single slice 3×3 voxel masks within which we calculated mean FA.

Probabilistic tractography

Probabilistic diffusion tractography was carried out according to previously described methods (Behrens et al., 2003b) using FMRIB's Diffusion Toolbox from within FSL. Using Bayesian techniques, we estimated a probability distribution function (pdf) on the principal fibre direction at each voxel. We then generated probability distributions of connectivity between seed and all other points by repeatedly sampling connected pathways through this pdf field. The effect of this procedure is to build a probability distribution on the location of the dominant connection from the seed voxel. The reproducibility of tract measures was compared using the following three different methods for defining seed points:

1. Single voxels manually defined on individual FA maps (see above for criteria) were used to generate unconstrained paths.
2. 3×3 voxel ROIs centred on manually defined individual voxels were used to generate unconstrained paths.

Table 2
Results for region of interest analysis of FA and MD

Diffusion measure	ROI	Mean across subjects	Variability measures		
			Inter-session		Inter-subject
			Mean inter-session CV (%)	Inter-subject CV—based on mean across sessions (%)	Inter-subject CV—based on session one only (%)
FA	CB-L	0.59	4.86	7.50	7.54
	CB-R	0.54	5.65	9.27	9.58
	OR-L	0.44	5.98	9.03	12.37
	OR-R	0.44	5.79	5.73	5.83
	PT-L	0.61	4.31	2.90	6.46
	PT-R	0.60	6.14	4.22	7.08
	Genu	0.67	4.81	9.41	12.63
MD		($\times 10^{-4}$)			
	CB-L	6.7	2.93	2.73	2.27
	CB-R	6.6	4.70	4.53	4.93
	OR-L	7.8	4.36	9.28	10.72
	OR-R	7.3	2.45	5.05	5.30
	PT-L	8.3	3.15	3.15	6.43
	PT-R	8.3	4.40	4.87	4.89
Genu	7.6	6.28	10.08	12.97	

Values for mean FA and MD within manually defined ROIs are given in column 3. Inter-session and inter-subject variability of mean FA and MD is given in columns 4–6. CB=cingulum bundle, OR=optic radiations, PT=pyramidal tract.

Table 3
Results of tractography analysis for mean tract FA

ROI	Mean FA across subjects			Variability measures								
				Inter-session			Inter-subject					
Space	DTI		MNI	Mean inter-session CV (%)			Inter-subject CV—based on mean across sessions (%)			Inter-subject CV—based on session one only (%)		
Seed method	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI
CB-L	0.27	0.25	0.35	8.58	6.42	3.18	13.08	7.50	7.31	18.19	13.35	8.32
CB-R	0.27	0.26	0.32	8.96	6.64	4.32	6.91	5.77	8.30	13.56	11.90	9.27
OR-L	0.29	0.30	0.45	8.75	5.12	2.12	3.24	8.04	5.41	7.48	8.39	6.26
OR-R	0.32	0.30	0.45	7.69	3.29	2.27	7.60	6.63	5.36	13.07	8.87	6.30
PT-L	0.36	0.34	0.45	4.09	3.92	1.52	4.84	4.21	3.20	5.93	6.87	3.30
PT-R	0.35	0.33	0.45	6.90	3.58	1.32	3.41	5.16	4.55	2.74	6.33	4.03
GENU	0.30	0.28	0.43	8.61	5.14	1.94	10.89	7.73	5.89	8.71	6.49	6.90

Columns 2–4 give values of mean FA from within tracts defined using 3 different seed methods. Remaining columns report variability of mean tract FA between sessions and between subjects for each seed method. All results are for data with 60 diffusion encoding directions. Single voxel and 3 × 3 voxel seeds are defined in DTI space. The 2 ROI seed method uses masks defined in standard (MNI) space. CB=cingulum bundle, OR=optic radiations, PT=pyramidal tract.

3. A two ROI approach using common standard space seed and target masks (see below for criteria).

These masks were used to generate estimated pathways from seed regions. In the resulting images, voxel values represented the probability of connection to the seed point. The connection probability images were thresholded at a connectivity value of 10 particles (5000 particles were seeded from each seed voxel) and binarised to define tract masks. (Note that we have chosen to use a fixed, arbitrary, low connectivity threshold, simply to exclude voxels with very low connectivity values. By using a fixed threshold for each tract, we will effectively have slightly different thresholding for the different seed methods, and for seed masks of difference sizes. Thresholding will, however, be comparable across sessions and subjects which is the main consideration here). The tract masks were then used to mask individual session FA and MD maps either in the space of the original diffusion images (for ROI methods 1 and 2 above) or after registration of the FA and MD maps to standard space (for ROI method 3) and mean FA and MD from within tracts was calculated.

For all measures, we calculated coefficients of variation (CV=standard deviation/mean × 100%). CV values were compared using repeated measures ANOVA within SPSS v. 11.

ROI definition: standard space

We defined regions of interest in standard (MNI) space (Evans et al., 2003) to constrain probabilistic tractography. Four types of mask were created:

1. Seed masks: Probabilistic tractography was seeded from all voxels within the seed mask.
2. Target masks: Only those pathways which reached the target mask were retained.
3. Termination masks: Masks adjacent to seed and target masks were used to terminate pathways beyond these regions.
4. Removal masks: Masks of certain regions (e.g., the opposite hemisphere) were used to remove (rather than just terminate) any pathways that entered these regions.

Table 4
Results of tractography analysis for mean tract MD

ROI	Mean tract MD across subjects (× 10 ⁻⁴)			Variability measures								
				Inter-session			Inter-subject					
Space	DTI		MNI	Mean inter-session CV (%)			Inter-subject CV—based on mean across sessions (%)			Inter-subject CV—based on session one only (%)		
Seed method	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI	Single voxel	3 × 3 voxel	2 ROI
CB-L	8.3	8.6	7.5	3.58	2.90	2.14	2.84	3.70	2.22	5.01	4.57	3.07
CB-R	8.3	8.7	7.5	4.12	2.75	1.61	2.74	3.40	2.82	4.09	3.63	2.76
OR-L	9.0	9.0	8.1	7.65	2.70	1.68	13.32	8.85	7.21	13.72	8.85	7.95
OR-R	9.0	8.6	8.0	3.75	2.72	1.49	3.62	3.51	4.27	6.02	2.54	5.32
PT-L	8.4	8.3	7.0	5.55	2.67	0.99	3.96	3.35	1.50	7.04	5.91	1.37
PT-R	8.3	8.5	7.1	2.95	3.57	1.24	7.23	3.18	1.63	5.22	2.47	1.78
GENU	9.0	9.1	8.2	4.30	2.34	1.86	7.88	5.86	4.73	7.67	5.06	4.92

Columns 2–4 give values of mean diffusivity (MD) within tracts defined using 3 different seed methods. Remaining columns report variability of mean MD between sessions and between subjects for each seed method. All results are for data with 60 diffusion encoding directions. Single voxel and 3 × 3 voxel seeds are defined in DTI space. The 2 ROI seed method uses masks defined in standard (MNI) space. CB=cingulum bundle, OR=optic radiations, PT=pyramidal tract.

Table 5
Results of tractography analysis for mean tract volume

ROI	Mean tract volume across subjects (voxels)			Variability measures								
				Inter-session			Inter-subject					
Space				Mean inter-session CV (%)			Inter-subject CV—based on mean across sessions (%)			Inter-subject CV—based on session one only (%)		
				DTI	MNI	2 ROI	DTI	MNI	2 ROI	DTI	MNI	2 ROI
Seed method	Single voxel	3×3 voxel	2 ROI	Single voxel	3×3 voxel	2 ROI	Single voxel	3×3 voxel	2 ROI	Single voxel	3×3 voxel	2 ROI
CB-L	6433	23,817	4212	18.47	11.36	8.51	15.02	15.51	28.90	25.14	17.74	31.73
CB-R	6195	23,758	4504	18.86	10.13	13.02	19.26	6.31	28.74	28.93	13.17	27.98
OR-L	10,372	45,583	4368	23.21	12.15	10.36	12.80	11.47	35.95	22.63	13.02	37.73
OR-R	9166	41,119	4036	20.42	12.54	9.44	23.87	20.78	28.54	35.35	25.05	33.49
PT-L	4873	39,426	7193	18.39	13.87	8.15	34.73	25.00	19.06	50.30	32.49	13.82
PT-R	5177	39,273	6861	30.58	10.58	7.59	42.89	19.91	30.42	70.27	23.56	32.03
GENU	4122	12,120	14,848	12.31	14.31	5.03	24.28	22.44	9.00	21.27	20.46	13.63

Columns 2–4 give values of tract volume for tracts defined using 3 different seed methods. Remaining columns report variability of mean tract volume between sessions and between subjects for each seed method. All results are for data with 60 diffusion encoding directions. Single voxel and 3×3 voxel seeds are defined in DTI space. The 2 ROI seed method uses 2 masks defined in standard (MNI) space. CB=cingulum bundle, OR=optic radiations, PT=pyramidal tract.

Standard space masks were created using the following criteria (Fig. 1):

1. Cingulum bundle (CB): The seed mask was located in the white matter of the cingulum bundle, above the body of the corpus callosum, on a single coronal slice at $Y=-2$, extending from $X=-4$ to $X=-12$ and from $Z=28$ to $Z=40$. The target mask included two lines of voxels through the white matter of the cingulum bundle—one located anteriorly, centred approximately at $Y=34$ and one located posteriorly, centred approximately at $Y=-42$. Termination masks located just in front of the anterior target, and just behind the posterior target, were used to terminate pathways passing beyond the targets. No removal mask was used.
2. Genu of CC (GCC): The seed mask was located at the genu of the corpus callosum on a single mid-sagittal slice at $X=0$. In an anterior posterior direction, it included all the corpus callosum and extended from $Z=-4$ to $Z=12$. The target mask included bilateral white matter of the prefrontal cortex, at $Y=42$, extending from $X=\pm 14$ to $X=\pm 28$ and from $Z=-6$ to $Z=22$. A termination mask of voxels surrounding the target mask plus a region of the same size as the target mask but one slice anterior, was used to terminate pathways travelling beyond the target. No removal mask was used.
3. Optic radiation (OR): The seed mask was a line of voxels through the white matter of the optic radiations close to the LGN, centred approximately at $Y=-28$, extending from $Z=-6$ to $Z=2$. The target mask was a line of voxels through the optic radiations as they approach visual cortex, centred approximately at $Y=-84$ and on the same axial slices as the seed mask. Termination masks were located just anterior to the seed mask, and just posterior to the target masks to cut off pathways beyond these points. Finally, a removal mask of the entire contralateral hemisphere and the ipsilateral hemisphere anterior to $Y=-6$ was used to exclude any paths that entered those regions.
4. Pyramidal tract (PT): The seed mask was a single axial slice of the whole cerebral peduncle at the level of $Z=-14$. The target mask was a section of white matter below the motor cortex, at the level

of $Z=52$, extending from $X=-6$ to $X=-34$ and from $Y=-28$ to $Y=-8$. A termination mask of the whole brain inferior to $Z=-18$ was used to cut off pathways extending below the peduncle. A termination mask including voxels surrounding the target mask at the same axial slice, plus an area the same size as the target mask, located one slice above the target mask, was used to terminate pathways travelling beyond the target mask. Finally, a removal mask of the opposite hemisphere was used to exclude any paths that entered those regions.

Comparing numbers of diffusion encoding directions

We acquired data with 60 diffusion-encoding directions. Acquisition of fewer diffusion directions is more typical, however, particularly with clinical studies where short scan times are desirable. To test whether the number of diffusion directions influences inter-session CV, we selected the 12 directions from our 60 which best satisfied the electrical repulsion model (Jones et al., 1999a) to create a 12 direction data set with which we generated tracts of interest using the same procedures described above. Note that these two data sets will differ not only in number of directions, but also in signal to noise, as the total number of measurements is greater for the 60 direction data (see Discussion).

Results

Histogram analysis

We derived histogram measures (mean, mode, peak height, standard deviation) for FA and MD within the brain white matter. Inter-session coefficients of variation (CVs) were less than 1.6% for all FA histogram measures (Table 1). Inter-subject CVs for FA ranged from 2.9 to 4.9%. Variability for MD histogram measures was higher: inter-session CVs were all under 3.6% while inter-subject CVs ranged from 1.6 to 20% (Table 1).

Region of interest FA and MD analyses

Single voxels of interest were manually defined within specific tracts on individual subject FA images (see methods). A single slice

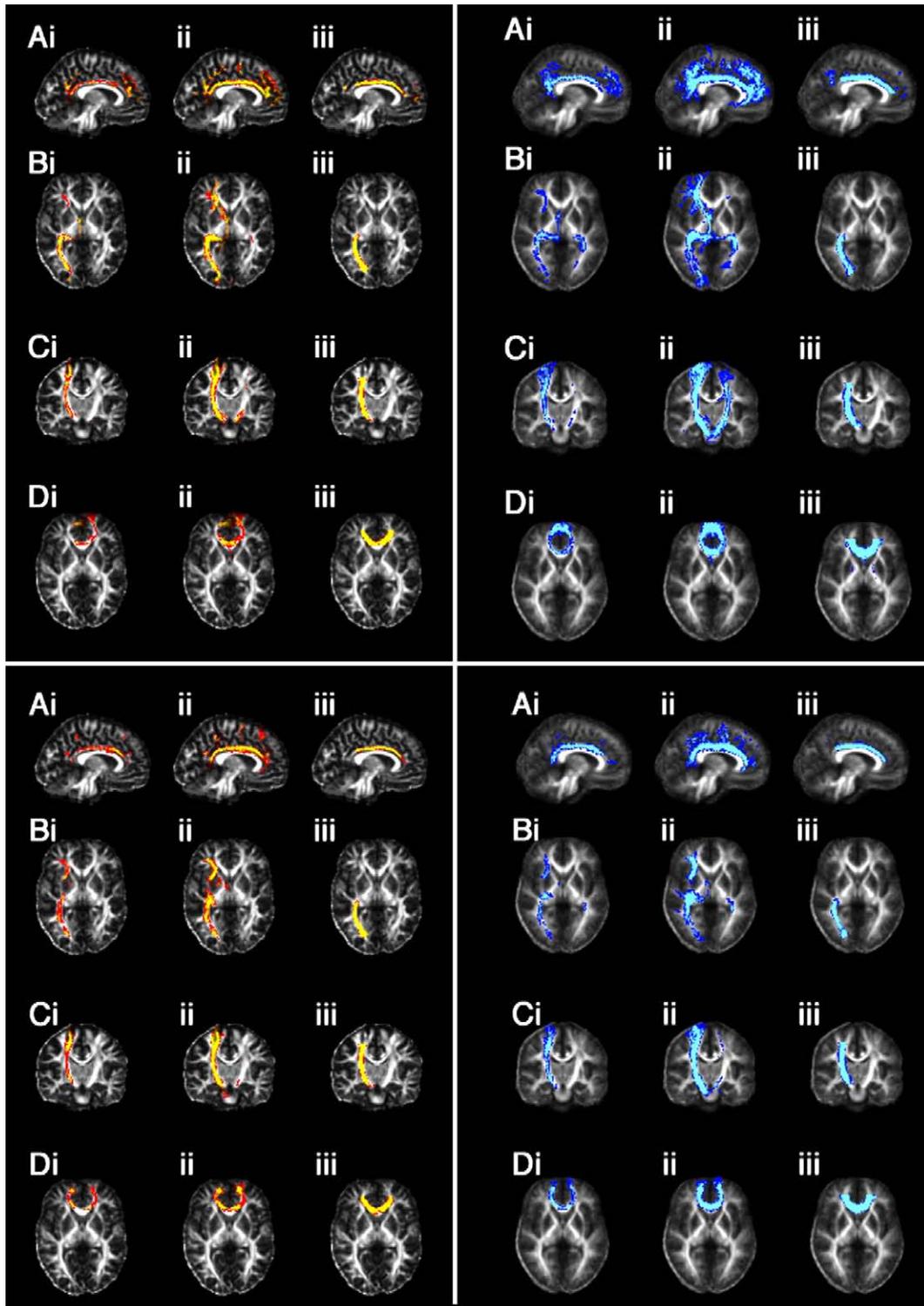


Fig. 2. Overlap of tracts between sessions (left) and between subjects (right) for data with 60 (top) or 12 (bottom) directions. Panel 1 (Top, left): Between session overlap of tracts for 60 direction data. Example shown for one subject. Background image is individual subject's FA map registered into standard space. Colour scale represents the number of sessions in which a voxel appeared in the tract from 1 (red) to 3 (yellow). Columns (i) to (iii) show overlap for different methods of seed definition: (i) single seed voxel; (ii) 3×3 voxel seed ROI; (iii) two ROIs in standard space. Each row shows results for a different tract. (A) Cingulum bundle, (B) optic radiation, (C) cortico-spinal tract, (D) genu of corpus callosum. Panel 2 (top, right): Between subject overlap of tracts for 60 direction data. Background image is group mean FA image. Colour scale represents the number of subjects in which a voxel appeared in the tract from 1 (blue) to 8 (turquoise). Columns (i) to (iii) show overlap for different methods of seed definition: (i) single seed voxel; (ii) 3×3 voxel seed ROI; (iii) two ROIs in standard space. Each row shows results for a different tract. (A) Cingulum bundle, (B) optic radiation, (C) cortico-spinal tract, (D) genu of corpus callosum. Panels 3 and 4 in the lower half of the figure show corresponding results for data with 12 diffusion encoding directions.

Table 6
Required sample size (per group) for between subject comparisons of mean tract FA

Tract	Cingulum bundle		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
Mean FA (SD)	0.35 (0.026)	0.32 (0.027)	0.45 (0.024)	0.45 (0.024)	0.45 (0.014)	0.45 (0.021)	0.43 (0.025)
Effect size	Required <i>N</i>						
2%	171	221	89	89	31	69	106
5%	28	36	15	15	7	13	18
10%	8	10	5	5	3	4	6
15%	4	5	3	3	–	3	3
20%	3	4	3	3	–	–	3

For each tract of interest, we report the mean and (between subject) standard deviation for mean FA and use these values to calculate the sample size required in each group to detect reductions in mean tract FA of 2–20% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

3×3 voxel mask was centred on these voxels of interest and the mean FA and MD values within the mask were found. Inter-session CVs for mean FA and MD defined in this way were in the order of 3–6% and inter-subject CVs were between 2 and 13% (Table 2).

Tractography analysis

Voxels of interest were used as seed points for probabilistic tractography. Resulting tracts were thresholded at a connectivity value of 10 and the mean FA (Table 3), MD (Table 4) and volume (Table 5) of the thresholded tract were found. The overlap of tracts across sessions and subjects can be seen in Fig. 2 (top half). To quantify this overlap for the most reproducible method of seed definition (standard space, two ROI—see below), we calculated the number of overlapping voxels for each pair of tracts and divided by the average tract volume. Across sessions we found mean overlap of 81% and across subjects we found mean overlap of 50%, reflecting individual variation in brain anatomy as well as any additional errors in registration across individuals.

Influences on inter-session reproducibility: DTI measures and method of seed definition

The effects of different factors on inter-session CV were tested using repeated measures ANOVAs. Our first model allowed for testing of hemispheric effects and so did not include results for the corpus callosum. We included factors of tract (CB, OR, PT), hemisphere (left, right), method of seed definition (manual single voxel, manual 3×3 voxel ROI, standard space two ROI) and measure (mean tract FA, mean tract MD, tract volume). We found that CVs calculated from mean tract MD (CV=3.2) or mean tract FA

(CV=4.9) were lower than those based on tract volume (CV=14.3) (main effect of measure $F=127.6$, $p<0.001$). The method of seed definition also influenced reproducibility with the lowest CV found for the two ROI standard space method (CV=4.5) followed by manually defined ROIs using a 3×3 voxel ROI (CV=6.5) then single seed voxel (CV=11.4) (main effect of method of seed definition, $F=81.4$, $p<0.001$). An interaction was found between measure and method of ROI definition ($F=13.5$, $p<0.001$) such that the effect of method of ROI definition on reproducibility is more marked when measuring tract volumes than when measuring mean diffusivity or mean FA. An interaction was also found between measure and tract ($F=3.8$, $p<0.02$) such that the improved reproducibility of mean tract diffusivity or FA compared to tract volume varied between tracts, being most prominent in PT, followed by OR, then CB.

Separate analyses for the left and right hemisphere, including data from the genu of the corpus callosum in each, confirmed the above results (main effect of measure, left $F=105.68$, $p<0.001$; right $F=91.9$, $p<0.001$; main effect of seed definition, left $F=41.22$, $p<0.001$; right $F=53.73$, $p<0.001$; interaction between measure and method of ROI definition, left $F=6.22$, $p<0.001$; right $F=5.45$, $p<0.001$).

Influences on inter-session reproducibility: number of diffusion directions

To test whether the number of diffusion directions influences inter-session CV, we created 12 direction data sets from our 60 direction data for comparison. The overlap in generated tracts across sessions and subjects for 12 direction data can be seen in the lower half of Fig. 2 and contrasted with the overlaps for the 60 direction data shown in the upper half of Fig. 2. The 12 direction

Table 7
Required sample size (per group) for between subject comparisons of mean tract MD

Tract	Cingulum bundle		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
MD (SD) × 10 ⁻⁴	7.5 (0.17)	7.5 (0.21)	8.1 (0.58)	8.0 (0.34)	7.0 (0.11)	7.1 (0.12)	8.1 (0.39)
Effect size	Required <i>N</i>						
2%	17	25	159	57	9	10	72
5%	4	5	26	10	3	3	13
10%	2	3	7	4	–	–	4
15%	–	2	4	3	–	–	3
20%	–	–	–	–	–	–	–

For each tract of interest, we report the mean and (between subject) standard deviation for mean MD and use these values to calculate the sample size required in each group to detect reductions in mean tract MD of 2–20% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

Table 8
Required sample size (per group) for between subject comparisons of tract volume

Tract	Cingulum		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
Mean volume (SD)	4212 (1217)	4504 (1295)	4368 (1570)	4036 (1152)	7193 (1370)	6861 (2087)	14848 (1336)
Effect size	Required <i>N</i>						
2%	2596	2561	4027	2502	1120	2870	251
5%	413	411	643	403	180	459	41
10%	105	104	161	102	46	116	11
15%	47	47	72	46	21	52	6
20%	27	27	41	26	12	30	4

For each tract of interest, we report the mean and (between subject) standard deviation for tract volume and use these values to calculate the sample size required in each group to detect reductions in tract volume of 2–20% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

data generally recover very similar pathways to those found with 60 direction data although there are instances where tractography in the 60 direction data is clearly more sensitive to branching of connectivity distributions, particularly approaching cortex; this difference is most apparent in the between-subject reproducibility maps and is explored quantitatively below.

We found the inter-session and inter-subject CVs for values of mean FA, mean MD and tract volume as above. We used a repeated measures ANOVA to compare inter-session CV from the 12 direction and 60 direction data sets including the factor of number of diffusion directions (12,60) in addition to the factors described above (tract, hemisphere, method of seed definition and measure). For simplicity, we report only those factors that include the factor number of directions. We did not find a main effect of number of diffusion directions but did find 3-way interactions between number of directions, tract and measure ($F=7.14$, $p<0.001$), and between number of directions, measure and method of seed definition ($F=5.5$, $p<0.02$) which are explored in more detail below.

To clarify the effects of number of diffusion directions, we ran follow-up ANOVAs to test separately each method of seed definition. The number of diffusion directions had a significant main effect on CV acquired using a single manually placed seed voxel ($F=22.6$, $p<0.003$) reflecting a lower inter-session CV for 12 direction data compared to 60 direction data. There were no significant main effects or interactions of direction on ANOVAs based on CVs from 3×3 ROI seed masks or the 2 ROI standard space method. It is important to note, however, that although CVs tended to be similar, or even smaller, in 12 direction compared to 60 direction data, it was not the case that the recovered tracts were the same for both data sets.

To explore further the effects of number of diffusion directions, we ran repeated measures ANOVA on the measurements themselves (i.e., mean tract FA, mean tract MD and tract volume). We

found that the tracts generated in 12 direction data were significantly smaller (main effect of number of directions on tract volume, $F=180.89$, $p<0.001$) and had significantly higher mean FA (main effect of direction on mean FA, $F=619.14$, $p<0.001$) and lower mean MD (main effect of direction on mean MD, $F=142.5$, $p<0.001$) than those generated in 60 direction data.

Implications for power and sample size calculations

The values of CV reported here should assist in estimation of required sample sizes to detect effects of a given size between groups of subjects or within-subjects, over time. Using estimates of variability between subjects, we have calculated the required sample size to detect a between-group difference of 2–20% in mean tract FA (Table 6), mean tract MD (Table 7) or tract volume (Table 8). Using estimates of variability between sessions, we have calculated required samples size to detect within-subject, between-session differences of 1–10% (Tables 9–11). All sample size calculations use the 60 direction data, and the two ROI standard space method for defining tracts, with a significance level of 0.05 and power of 0.8 (note that calculations of required sample sizes for 12 direction data, or for other methods of seed definition, can be made using the data provided in the tables above). Samples sizes were calculated using Russell Lenth's Java applets for Power and Sample Size (www.stat.uiowa.edu/~rlenth/). The significantly reduced variability between sessions or subjects found for mean tract FA and MD compared to tract volumes means that much lower sample sizes are required to detect changes of any given size in mean tract FA or MD. Note that, for the examples given here, significance for one-tailed tests has been computed, assuming that there is a hypothesis for directional changes in FA, MD or tract volume; two-tailed significance would require greater subject numbers. Note also that the varying reproducibility across tracts

Table 9
Required sample size for between session (within subject) comparisons of mean tract FA

Tract	Cingulum bundle		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
Mean FA (SD)	0.35 (0.011)	0.32 (0.014)	0.45 (0.009)	0.45 (0.010)	0.45 (0.007)	0.45 (0.006)	0.43 (0.008)
Effect size	Required <i>N</i>						
1%	62	120	26	32	17	13	23
2%	17	31	8	9	5	4	7
5%	5	7	3	3	3	3	3

For each tract of interest, we report the mean and (between session) standard deviation for mean tract FA and use these values to calculate the sample size required in each group to detect reductions in mean tract FA over time of 1–5% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

Table 10
Required sample size for between session (within subject) comparisons of mean tract MD

Tract	Cingulum bundle		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
Mean MD (SD)	7.5 (0.16)	7.5 (0.12)	8.1 (0.14)	8.0 (0.12)	7.0 (0.07)	7.1 (0.09)	8.1 (0.15)
Effect size	Required <i>N</i>						
1%	30	18	20	16	8	12	23
2%	9	6	7	6	4	5	7
5%	4	3	3	3	3	3	3

For each tract of interest, we report the mean and (between session) standard deviation for mean tract MD and use these values to calculate the sample size required in each group to detect reductions in mean tract MD over time of 1–5% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

means that the numbers here cannot be generalised in a straightforward way to other pathways.

Discussion

Quantification of the reproducibility of measures based on diffusion tensor imaging and tractography is a prerequisite for the design of clinical studies. Here, we quantified the reproducibility of histogram and ROI-based DTI measures and also of FA, MD and tract volume measurements along pathways defined using probabilistic tractography.

Histogram measures of FA produced very low coefficients of variation, both between session (below 2%) and between subjects (3–5%). For MD, CVs between session were all below 5% and between subject ranged from 1 to 20%. Whole brain white matter histogram measures, however, do not allow localisation of changes to specific regions of interest. Mean FA or MD taken from within manually placed regions of interest was slightly more variable, though still within acceptable ranges (between session CV=2–6%, between subject CV=3–12%).

Hypotheses of white matter changes in disease will often concern specific pathways. It is therefore useful to extract quantitative measures from along pathways defined in each individual using tractography (Ciccarelli et al., 2006; Ding et al., 2003; Gong et al., 2005; Jones et al., 2005; Kanaan et al., 2006). We calculated the reproducibility of measures derived in this way, using a number of different techniques for defining a tract of interest. Measures of FA or MD taken from along tracts were more reproducible than tract volumes. The method of seed definition also influenced reproducibility, with tracts defined using a two-ROI method with standard space masks being more reproducible than those defined using seeds placed manually on

DTI images. For mean FA taken from tracts defined using the two ROI approach, inter-session CVs were all below 5% and inter-subject CVs were all below 10%; for mean MD inter-session CVs were all below 3% and inter-subject CVs were below 8%. These values are in broad agreement with previously reported figures on a smaller group of subjects scanned twice (Ciccarelli et al., 2003a).

Note that the advantages associated with defining seeds using the two-ROI method with standard space masks could be due to many different factors—for example, only seed ROIs (rather than additional target and exclusion ROIs) were used for the manual methods; also, the shape of our manual ROIs was fixed across all tracts (as a single voxel, or a 3×3 square) whereas the shapes of our standard space seeds were allowed to vary across tracts. Additionally, the manual method could suffer additional variability of ROI placement although previous studies suggest that voxel placement differs by less than a voxel with clear criteria such as those used here (Ciccarelli et al., 2003a).

The majority of our reported findings are based on data with 60 diffusion encoding directions. However, such data take time to acquire, and clinical studies typically acquire fewer diffusion directions. We therefore explored the effects of number of diffusion directions on tractography. We found no significant main effects of number of diffusion directions on inter-session coefficients of variation. Tracts derived using 60 direction data, however, had a consistently greater volume than those found using 12 direction data. This fits with qualitative inspection (compare upper and lower panels in the right hand side of Fig. 2, for example), i.e., tracts generated in 60 direction data tend to travel further, and into lower FA regions (e.g., approaching grey matter), whereas tracts generated in 12 direction data tend to be restricted to the major fibre bundles. Comparison of CV values

Table 11
Required sample size for between session (within subject) comparisons of tract volume

Tract	Cingulum		Optic radiation		Pyramidal tract		Genu of CC
	Left	Right	Left	Right	Left	Right	
Mean volume (SD)	4212 (339)	4504 (565)	4368 (445)	4036 (371)	7193 (657)	6861 (490)	14,848 (762)
Effect size	Required <i>N</i>						
1%	404	976	634	533	516	313	165
2%	102	245	159	134	130	79	42
5%	17	40	27	22	22	14	8
10%	5	11	8	7	7	5	3

For each tract of interest, we report the mean and (between session) standard deviation for tract volume and use these values to calculate the sample size required in each group to detect reductions in tract volume over time of 1–5% with a one-tailed significance level of 0.05, power of 0.8, equal sample sizes.

alone does not fully reflect the differences between the two data sets. For this reason, the differences between 60 and 12 direction tractography are minimal when large bundles, such as the corpus callosum, are tracked, but are more apparent in smaller or more tortuous paths, such as the optic radiations (Fig. 2). In the comparison made here, differences between 60 and 12 direction data could be due to the lower signal to noise of the 12 direction data set (fewer data points) or to the lower angular resolution in the 12 direction data.

In assessing the reproducibility of the different approaches reported here, it is important to keep in mind the need to balance reproducibility with sensitivity in clinical studies. A measure (such as the mean of whole brain white matter FA histogram) may have a very low coefficient of variation, but might be insensitive to change between subjects or over time. Likewise, a measure with relatively poor reproducibility (such as tract volume), that is sensitive therefore to subtle variations in data from session to session (whatever their source), may also show greater sensitivity to clinically relevant change—though detecting clinically relevant changes over and above changes due to noise will clearly be challenging.

Power analyses are a critical step in the design of clinical studies to look for alterations in white matter structure between groups of subjects or over time. The values provided here allow for calculation of required sample sizes for given effect sizes. Some examples for changes in FA, MD or tract volume are given in Tables 6–11 above. Due to the increased variability in volume measurements, required sample sizes are much smaller for detection of changes of a given size in FA or MD, rather than tract volume. Furthermore, as inter-session variability is much lower than inter-subject variability, studies designed to look for within-subject changes over time should be able to detect small (e.g., 2%) changes in FA with reasonable subject numbers (e.g., 5–30 subjects, depending on tract of interest).

For clinical studies of white matter integrity, choice of measure and tract will of course depend on the type and site of pathology in question. Measures tended to be most reproducible for the pyramidal tracts and most variable for the cingulum bundle. The pyramidal tracts fall within large white matter fibre bundles and are very easy to trace, whereas the cingulum bundle is a smaller structure, sensitive to partial volume effects from adjacent grey matter. The difference in reproducibility across the pathways highlights the need for power analyses to be performed based on data from the pathways of interest.

Acknowledgment

We are grateful to the Medical Research Council, UK for funding (HJ-B, TEJB).

References

- Abe, O., Yamada, H., Masutani, Y., Aoki, S., Kunimatsu, A., Yamasue, H., et al., 2004. Amyotrophic lateral sclerosis: diffusion tensor tractography and voxel-based analysis. *NMR Biomed.* 17, 411–416.
- Aoki, S., Iwata, N.K., Masutani, Y., Yoshida, M., Abe, O., Ugawa, Y., et al., 2005. Quantitative evaluation of the pyramidal tract segmented by diffusion tensor tractography: feasibility study in patients with amyotrophic lateral sclerosis. *Radiat. Med.* 23, 195–199.
- Bammer, R., Augustin, M., Strasser-Fuchs, S., Seifert, T., Kapeller, P., Stollberger, R., et al., 2000. Magnetic resonance diffusion tensor imaging for characterizing diffuse and focal white matter abnormalities in multiple sclerosis. *Magn. Reson. Med.* 44, 583–591.
- Basser, P.J., Mattiello, J., LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. *Biophys. J.* 66, 259.
- Beaulieu, C., 2002. The basis of anisotropic water diffusion in the nervous system—A technical review. *NMR Biomed.* 15, 435.
- Behrens, T.E.J., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., et al., 2003a. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat. Neurosci.* 6, 750.
- Behrens, T.E.J., Woolrich, M.W., Jenkinson, M., Johansen-Berg, H., Nunes, R.G., Clare, S., et al., 2003b. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magn. Reson. Med.* 50, 1077.
- Bozzali, M., Franceschi, M., Falini, A., Pontesilli, S., Cercignani, M., Magnani, G., et al., 2001. Quantification of tissue damage in AD using diffusion tensor and magnetization transfer MRI. *Neurology* 57, 1135–1137.
- Buchsbaum, M.S., Tang, C.Y., Peled, S., Gudbjartsson, H., Lu, D., Hazlett, E.A., et al., 1998. MRI white matter diffusion anisotropy and PET metabolic rate in schizophrenia. *NeuroReport* 9, 425.
- Cercignani, M., Bammer, R., Sormani, M.P., Fazekas, F., Filippi, M., 2003. Inter-sequence and inter-imaging unit variability of diffusion tensor MR imaging histogram-derived metrics of the brain in healthy volunteers. *Am. J. Neuroradiol.* 24, 638–643.
- Chabriat, H., Pappata, S., Poupon, C., Clark, C.A., Vahedi, K., Poupon, F., et al., 1999. Clinical severity in CADASIL related to ultrastructural damage in white matter: in vivo study with diffusion tensor MRI. *Stroke* 30, 2637.
- Charlton, R.A., Barrick, T.R., McIntyre, D.J., Shen, Y., O'Sullivan, M., Howe, F.A., et al., 2006. White matter damage on diffusion tensor imaging correlates with age-related cognitive decline. *Neurology* 66, 217–222.
- Ciccarelli, O., Parker, G.J., Toosy, A.T., Wheeler-Kingshott, C.A., Barker, G.J., Boulby, P.A., et al., 2003a. From diffusion tractography to quantitative white matter tract measures: a reproducibility study. *NeuroImage* 18, 348.
- Ciccarelli, O., Werring, D.J., Barker, G.J., Griffin, C.M., Wheeler-Kingshott, C.A., Miller, D.H., et al., 2003b. A study of the mechanisms of normal-appearing white matter damage in multiple sclerosis using diffusion tensor imaging—evidence of Wallerian degeneration. *J. Neurol.* 250, 287–292.
- Ciccarelli, O., Behrens, T.E., Altmann, D.R., Orrell, R.W., Howard, R.S., Johansen-Berg, H., et al., 2006. Probabilistic diffusion tractography: a potential tool to assess the rate of disease progression in ALS. *Brain* 129, 1859–1871.
- Ding, Z., Gore, J.C., Anderson, A.W., 2003. Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magn. Reson. Med.* 49, 716–721.
- Evans, A.C., Collins, D.L., Mills, S.R., Brown, E.D., Kelly, R.L., Peters, T.M., 2003. 3D statistical neuroanatomical models from 305 MRI volumes. *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.* 95, 1813.
- Foong, J., Maier, M., Clark, C.A., Barker, G.J., Miller, D.H., Ron, M.A., 2000. Neuropathological abnormalities of the corpus callosum in schizophrenia: a diffusion tensor imaging study. *J. Neurol., Neurosurg. Psychiatry* 68, 242–244.
- Gong, G., Jiang, T., Zhu, C., Zang, Y., He, Y., Xie, S., et al., 2005. Side and handedness effects on the cingulum from diffusion tensor imaging. *NeuroReport* 16, 1701–1705.
- Griffin, C.M., Chard, D.T., Ciccarelli, O., Kapoor, B., Barker, G.J., Thompson, A.I., et al., 2001. Diffusion tensor imaging in early relapsing—remitting multiple sclerosis. *Mult. Scler.* 7, 290–297.
- Huang, H., Zhang, J., Jiang, H., Wakana, S., Poetscher, L., Miller, M.I., et al., 2005. DTI tractography based parcellation of white matter: application to the mid-sagittal morphology of corpus callosum. *NeuroImage* 26, 195–205.

- Jenkinson, M., Smith, S.M., 2001. Global optimisation for robust affine registration. *Med. Image Anal.* 5, 143.
- Jones, D.K., Horsfield, M.A., Simmons, A., 1999a. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn. Reson. Med.* 42, 515.
- Jones, D.K., Simmons, A., Williams, S.C., Horsfield, M.A., 1999b. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn. Reson. Med.* 42, 37.
- Jones, D.K., Catani, M., Pierpaoli, C., Reeves, S.J., Shergill, S.S., O'Sullivan, M., et al., 2005. Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. *Hum. Brain Mapp.* 27, 230–238.
- Kanaan, R.A., Shergill, S.S., Barker, G.J., Catani, M., Ng, V.W., Howard, R., et al., 2006. Tract-specific anisotropy measurements in diffusion tensor imaging. *Psychiatry Res.* 146, 73–82.
- Lazar, M., Alexander, A.L., 2005. Bootstrap white matter tractography (BOOT-TRAC). *NeuroImage* 24, 524–532.
- McKinstry, R.C., Mathur, A., Miller, J.H., Ozcan, A., Snyder, A.Z., Schefft, G.L., et al., 2002. Radial organization of developing preterm human cerebral cortex revealed by non-invasive water diffusion anisotropy MRI. *Cereb. Cortex* 12, 1237–1243.
- Molko, N., Pappata, S., Mangin, J.F., Poupon, F., LeBihan, D., Bousser, M.G., et al., 2002. Monitoring disease progression in CADASIL with diffusion magnetic resonance imaging: a study with whole brain histogram analysis. *Stroke* 33, 2902–2908.
- Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C., 1999. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann. Neurol.* 45, 265.
- O'Sullivan, M., Jones, D.K., Summers, P.E., Morris, R.G., Williams, S.C., Markus, H.S., 2001. Evidence for cortical “disconnection” as a mechanism of age-related cognitive decline. *Neurology* 57, 632–638.
- Parker, G.J., Luzzi, S., Alexander, D.C., Wheeler-Kingshott, C.A., Ciccarelli, O., Lambon Ralph, M.A., 2005. Lateralization of ventral and dorsal auditory-language pathways in the human brain. *NeuroImage* 24, 656–666.
- Pfefferbaum, A., Sullivan, E.V., Hedehus, M., Lim, K.O., Adalsteinsson, E., Moseley, M., 2000. Age-related decline in brain white matter anisotropy measured with spatially corrected echo-planar diffusion tensor imaging. *Magn. Reson. Med.* 44, 259.
- Pfefferbaum, A., Adalsteinsson, E., Sullivan, E.V., 2003. Replicability of diffusion tensor imaging measurements of fractional anisotropy and trace in brain. *J. Magn. Reson. Imaging* 18, 427–433.
- Rocca, M.A., Colombo, B., Inglese, M., Codella, M., Comi, G., Filippi, M., 2003. A diffusion tensor magnetic resonance imaging study of brain tissue from patients with migraine. *J. Neurol., Neurosurg. Psychiatry* 74, 501–503.
- Rose, S.E., Chen, F., Chalk, J.B., Zelaya, F.O., Strugnell, W.E., Benson, M., et al., 2000. Loss of connectivity in Alzheimer's disease: an evaluation of white matter tract integrity with colour coded MR diffusion tensor imaging. *J. Neurol., Neurosurg. Psychiatry* 69, 528–530.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., et al., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23 (Suppl. 1), S208–S219.
- Toosy, A.T., Werring, D.J., Orrell, R.W., Howard, R.S., King, M.D., Barker, G.J., et al., 2003. Diffusion tensor imaging detects cortico-spinal tract involvement at multiple levels in amyotrophic lateral sclerosis. *J. Neurol., Neurosurg. Psychiatry* 74, 1250.
- Toosy, A.T., Ciccarelli, O., Parker, G.J., Wheeler-Kingshott, C.A., Miller, D.H., Thompson, A.J., 2004. Characterizing function–structure relationships in the human visual system with functional MRI and diffusion tensor imaging. *NeuroImage* 21, 1452–1463.
- Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation–maximization algorithm. *IEEE Trans. Med. Imaging* 20, 45.