

Published in final edited form as:

Neuroimage. 2007 July 1; 36(3): 630–644.

Reproducibility of Quantitative Tractography Methods Applied to Cerebral White Matter

Setsu Wakana^{1,2}, Arvind Caprihan³, Martina M. Panzenboeck⁴, James H. Fallon⁴, Michele Perry⁵, Randy L. Gollub⁶, Kegang Hua¹, Jiayang Zhang¹, Hangyi Jiang^{1,2}, Prachi Dubey¹, Ari Blitz², Peter van Zijl^{1,2}, and Susumu Mori^{1,2}

1) F. M. Kirby center for functional brain imaging, Kennedy Krieger Institute Baltimore, Maryland *2) Department of Radiology, Johns Hopkins University, School of Medicine, Baltimore, Maryland* *3) New Mexico Resonance and The MIND Institute, , Albuquerque New, Mexico* *4) Department of Anatomy and Neurobiology, University of California Irvine, Irvine, California* *5) Department of Psychiatry, University of California, San Diego, California* *6) Department of Psychiatry, Massachusetts General Hospital, Charlestown, Massachusetts*

Abstract

Tractography based on diffusion tensor imaging (DTI) allows visualization of white matter tracts. In this study, protocols to reconstruct eleven major white matter tracts are described. The protocols were refined by several iterations of intra- and inter-rater measurements and identification of sources of variability. Reproducibility of the established protocols was then tested by raters who did not have previous experience in tractography. The protocols were applied to a DTI database of adult normal subjects to study size, fractional anisotropy (FA), and T_2 of individual white matter tracts. Distinctive features in FA and T_2 were found for the corticospinal tract and callosal fibers. Hemispheric asymmetry was observed for the size of white matter tracts projecting to the temporal lobe. This protocol provides guidelines for reproducible DTI-based tract-specific quantification.

Keywords

Diffusion Tensor Imaging; Tractography; reproducibility; white matter

Introduction

Three-dimensional tract reconstruction (tractography) based on diffusion tensor imaging (DTI) is becoming a widely-used tool to study human white matter anatomy (Basser et al., 2000; Conturo et al., 1999; Jones et al., 1999b; Lazar et al., 2003; Mori et al., 1999; Mori et al., 2005; Parker et al., 2002; Poupon et al., 2000). This technology allows us to visualize trajectories of specific white matter fiber bundles and has potential to perform quantitative evaluation of properties of individual tracts. This provides exciting opportunities to assess the impact of diseases on specific white matter tracts. Once the location of a tract is defined, its size can be measured. Further, by superimposing tract coordinates on various MR parameter maps such as T_1 , T_2 , magnetization transfer ratio (MTR), and diffusion anisotropy, the

Correspondence to; Susumu Mori, Ph.D Department of Radiology, Johns Hopkins University, School of Medicine, 217 Traylor Bldg, 720 Rutland Ave. Baltimore, MD, 21205, USA Phone ;410-614-2702 FAX; 410-614-1948 E-mail; susumu@mri.jhu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

myelination and axonal status of individual tracts may be monitored (Glenn et al., 2003; Pagani et al., 2005; Partridge et al., 2004; Stieltjes et al., 2001; Virta et al., 1999; Wilson et al., 2003; Xue et al., 1999). However, questions remain whether the tractography results reflect true neuroanatomical details and are sufficiently reproducible to be used as a tool for quantitative analyses.

In terms of validity, there is mounting evidence that tracking results of many prominent white matter tracts agree very well with classical definitions based on postmortem studies (Basser et al., 2000; Catani et al., 2002; Conturo et al., 1999; Jellison et al., 2004; Mori et al., 1999; Mori et al., 2002; Poupon et al., 2000; Stieltjes et al., 2001; Wakana et al., 2004). On the other hand, it is well-known that the technique can produce false positive and false negative results due to noise, partial volume effects, and complex fiber architectures within a pixel (Pierpaoli et al., 2001; Wiegell et al., 2000). One way to raise the confidence level of validity is to employ anatomical constraints by employing multiple regions of interest (ROIs) (Conturo et al., 1999; Huang et al., 2004). This technique requires *a priori* knowledge about the trajectory and can be used only for well-characterized white matter tracts. This approach should reduce the number of false positive, but is unlikely to be 100% accurate.

While it is difficult to completely characterize the validity of tractography, we can measure its reproducibility. If we can develop a scheme to reproducibly define coordinates of specific white matter tracts, this technique should be a valuable tool to test hypotheses as to whether any of these specific tracts are involved in the disease of interest. Within a given set of imaging parameters, we expect reconstruction reproducibility to vary among white matter tracts depending on their size and trajectory. Therefore, it is important to establish reproducible tracking protocols.

One of the major sources of variability, in addition to noise and partial volume effects, comes from placement of reference ROIs to identify specific white matter tracts. By devising clever ROI drawing schemes, which are based on anatomical features of individual tracts and using regions that are sufficiently large, it is possible to come up with a robust protocol that is rather insensitive to small variations in the ROI drawing (Huang et al., 2004). The protocol can be iteratively improved by measuring intra and inter-rater variability and by identifying and describing the source of variations. The purpose of this paper is to develop such robust protocols and identify white matter tracts that can be reconstructed reproducibly. The reproducibility was characterized by spatial matching among different trials by the same rater (intra-rater) and multiple raters (inter-rater) using the same subject (inter-measurement). In the second part of the study we applied the established protocols to measure size, fractional anisotropy, and T_2 of each tract using our normal DTI database. This study leads to multi-parametric mapping of the normal white matter and the range of normal variations in a tract-by-tract basis.

Methods

Subjects

The study was approved by the institutional review board and informed consents including a HIPAA compliant data sharing agreement were obtained from all subjects. Ten healthy adults (mean 26.1 \pm 5.48 years old, male 5, female 5) participated. All subjects were free of current and past medical or neurological disorders. The raw and processed image data are accessible through our websites (lbam.med.jhmi.edu, godzilla.kennedykrieger.org and www.nbirn.net).

Imaging

A 1.5T MR unit (Gyrosan NT, Philips Medical Systems) was used. DTI data were acquired using single-shot echo-planar imaging with sensitivity encoding (SENSE, parallel-imaging

factor, 2.5) (Pruessmann et al., 1999). The imaging matrix was 96×96 with field of view of 240×240 mm (nominal resolution, 2.5mm), zero-filled to 256×256 pixels. Transverse sections of 2.5 mm thickness were acquired parallel to the anterior commissure/posterior commissure line. A total of 50-55 sections covered the entire cerebrum and brainstem without gaps. Diffusion weighting was encoded along 30 independent orientations (Jones et al., 1999a) using a b-value of $700 \text{ mm}^2/\text{s}$. Five additional images with minimal diffusion weighting ($b = 33 \text{ mm}^2/\text{sec}$) were also acquired. The scanning time per dataset was approximately 6 minutes including 2 minutes image reconstruction delay. To enhance the signal-to-noise ratio, imaging was repeated 3 times. Co-registered Magnetization-Prepared Rapid Gradient Echo (MPRAGE) images of the same resolution were recorded for anatomical guidance.

Single-shot-EPI, T2-weighted imaging (TR 4000msec, TEs 40 msec and 100 msec; image resolution and SENSE factor equal to DTI) was performed for T2 quantification. To ensure the best registration between DTI and T₂-weighted images, the same number of echoes was acquired after each excitation.

Data processing

The DTI datasets were transferred to a PC with windows platform and processed using the analysis software DTIstudio developed and distributed by this laboratory (H. Jiang & S. Mori, Johns Hopkins University and Kennedy Krieger institute, <http://godzilla.kennedykrieger.org> or <http://lbam.med.jhmi.edu>). Images were first realigned using the AIR program (Woods et al., 1998), in order to remove any potential small bulk motions that occurred during the scans. Subsequently, all diffusion-weighted images were visually inspected for apparent artifacts due to subject motion and instrumental malfunction by authors. The six elements of the diffusion tensor were calculated for each pixel using multi-variate linear fitting. After tensor diagonalization, three eigenvalues and eigenvectors were obtained and fractional anisotropy (FA) maps were calculated. The eigenvector associated with the largest eigenvalue was used as an indicator for fiber orientation. In the DTI color maps, red, green, and blue colors were assigned to right-left, anterior-posterior, and superior-inferior orientations, respectively. T₂ maps were calculated from two separated images with echo times of 40 ms and 100 ms.

Fiber Tracking and ROI drawing strategy

In this study, we investigated the reliability of reconstruction of white matter tracts in each hemisphere. For 3D tract reconstruction, the fiber assignment by continuous tracking or FACT method (Mori et al., 1999; Xue et al., 1999) was used with a fractional anisotropy threshold 0.2 and an inner product threshold of 0.75, which prohibited angles larger than 41 degrees during tracking. The fiber tracking was performed by DTIstudio. A multi-ROI approach was used to reconstruct tracts of interest which exploits existing anatomical knowledge of tract trajectories. Tracking was performed from all pixels inside the brain (brute-force approach) (Conturo et al., 1999; Huang et al., 2004) and results penetrating the manually defined ROIs were assigned to the specific tracts associated with the ROIs. When multiple ROIs were used for a tract of interest, we employed three types of operations, namely “AND”, “CUT”, and “NOT” (Fig. 1), the choice of which depends on the characteristic trajectory of each tract.

In our protocols, two sets of tracking results were generated; one using the “AND” operation (Fig. 1B) and the other using the “CUT” operation (Fig. 1D). For both cases, tracking results that penetrate two ROIs are extracted. When we use the “AND” operation (Fig. 1B, 1E), the results include three sections with different properties; one encompassing the tract before the first ROI (i.e. to the left side in figure 1), one between the two ROIs, and one encompassing the tract after the second ROI (i.e. to the left side in figure 1). Pathways in the middle section are constrained by two destinations with *a priori* knowledge while the other two sections are not. By using the “CUT” operation, only the middle section is retained (Fig. 1F). Use of the

“NOT” operation was sometimes necessary to remove a subset of unwanted projections from a reconstruction result. For identification of anatomical landmarks, color coded maps were used, but certain landmarks, such as sulci, could be more easily identified in co-registered images such as the anisotropy maps and diffusion-weighted images as will be discussed later. In this study, tracts were reconstructed as faithfully as possible based on classical *a priori* anatomical knowledge from postmortem studies using ROIs that were strategically located based on known fiber trajectories.

Reconstruction protocol

Tract #1: Cingulum cingulate gyrus part (Fig. 2)—The cingulum is defined as two separate segments; the upper part along the cingulate gyrus (CGC: cingulum cingulate gyrus part) and lower segment (Tract #2) along the ventral side of the hippocampus (CGH: cingulum hippocampal part).

For CGC (Fig.2), a coronal plane is selected at the middle of splenium of the corpus callosum (CC) using the mid-sagittal plane (Fig.2b) and a ROI as shown in Fig. 2a is drawn. For the second ROI, a coronal plane in the middle of genu of CC is selected using the mid-sagittal plane (Fig. 2d) and a second ROI is drawn to include the cingulum (Fig. 2c). The size of the second ROI doesn't affect the result as long as only the labeled cingulum is included.

Tract #2 ; Cingulum hippocampal part (Fig. 3)—The inferior segment of the cingulum runs along the ventral aspect of the hippocampus. For the first ROI, a coronal plane in the middle of the splenium of corpus callosum is selected using the mid-sagittal plane (Fig. 3b) and the cingulum below the corpus callosum is delineated. For the second ROI, a coronal slice anterior to the pons is selected using the mid-sagittal plane (Fig. 3d). The second ROI includes the cingulum which is already labeled by the tracking (Fig. 2c).

Tract #3: Cortico-spinal tract (CST) (Fig. 4)—In this protocol, we define the CST between the primary motor cortex and the midbrain. The first ROI defines the entire cerebral peduncle in an axial plane at the level of the decussation of the superior cerebellar peduncle (Fig. 4a), indicated by the arrow. By inspecting the reconstruction result from the first ROI, a bundle of trajectories that reach the primary motor cortex and the location of the central sulcus can be identified. The most ventral axial slice that can clearly identify the cleavage of the central sulcus in the tracking result (Fig. 4b) is selected and the bundle in the primary motor cortex is defined (Figs. 4c and d). As long as only the trajectories to the primary motor cortex are defined, the size of the second ROI can be arbitrary. The trajectories outside the two ROIs may cross the midline *via* the pontine crossing fibers and re-enter the contralateral hemisphere, which interferes with subsequent quantification procedures. These tracts should be removed by using the “NOT” operation across the entire mid-sagittal slice. Please note that if the “CUT” operation is used, which retains only coordinates between the two ROIs, this extra editing procedure using the “NOT” operation is unnecessary.

The tracking method described in this protocol usually reconstructs only the CST projecting to the medial cortical regions. The CST projection to the lateral areas of motor cortex can not be accurately reconstructed because there is a significant mixing of fibers with different orientations within the pixels as the CST passes through the massive bundle of association fibers.)

Tract #4: Anterior Thalamic Radiation (ATR) (Fig. 5)—A coronal slice is selected at the middle of the genu of corpus callosum (Fig. 5a, b). The first ROI defines the anterior limb of internal capsule (Fig.5 a). For the second ROI, a coronal plane at the anterior edge of pons (Fig.5 d) is selected and the entire thalamus is delineated (Fig.5 c). The trajectories outside the

two ROIs may cross the corpus callosum and enter the contralateral hemisphere. These trajectories are removed by a “NOT” operation across the entire mid-sagittal slice.

Tract #5: Superior Longitudinal Fasciculus (SLF) (Fig. 6)—We tested two different protocols to define the SLF. One is used to reconstruct the SLF as comprehensively as possible and the other is selective for isolating the trajectories to the temporal lobe (named SLFt in this paper).

For the 1st ROI, the lowest axial level in which the fornix can be identified is selected as a single intense structure (Fig. 6a, the fornix is indicated by an arrowhead). Then a coronal slice is selected at the middle of the posterior limb of internal capsule (Fig. 6a; white line). The core of the SLF can be identified as an intense green tract with a triangular shape (Fig. 6b, white triangle). The first ROI includes this core and all branches coming out from the triangular area (Fig. 6b and 6c).

For the second ROI, a coronal slice is selected at the middle of the splenium of corpus callosum using the mid-sagittal level (Fig. 6e and 6f). The second ROI includes all labeled fibers (Fig. 6e).

Tract #6: The temporal component of the SLF (Fig.6)—For the temporal component of SLF (SLFt), the first ROI is same as the body of SLF. For the second ROI, an axial slice is selected at the level of the anterior commissure (AC in Fig. 6g) and the projections located laterally to the sagittal stratum (SS in Fig. 6g) are delineated by the second ROI.

Tract #7: Inferior longitudinal fasciculus (ILF) (Fig.7)—Using a parasagittal slice, the posterior edge of the cingulum is identified (Fig. 7b, indicated by an arrow). A coronal slice is selected at that edge (coronal level #1 in Fig. 7b). The first ROI includes the entire hemisphere (Fig. 7c). For second ROI, the most posterior coronal slice in which the temporal lobe is not connected to the frontal lobe is selected (Fig. 7e). The second ROI includes the entire temporal lobe.

Tract #8: Inferior fronto-occipital fasciculus (IFO) (Fig 8)—For the first ROI, a coronal slice (the coronal level #2 in Fig.7b) is identified at the middle point between the posterior edge of the cingulum (slice level #1 in Fig. 7b) and the posterior edge of the parieto-occipital sulcus (slice level #3 in Fig. 7b). Although a color map is shown in Fig. 7b, another anatomical image such as the least-diffusion-weighted image better defines the parieto-occipital sulcus. The first ROI delineates the occipital lobe. The boundary between the occipital and parietal lobes is defined by linearly extrapolating the parieto-occipital sulcus (POS) to the lateral region (Fig. 8a). For the second ROI, a coronal slice is selected at the anterior edge of the genu of corpus callosum (Fig. 8c) and the entire hemisphere is delineated. These two ROIs are sometimes shared by the cingulum and some fibers relayed at the thalamus. These fibers clearly don't belong to the IFO and should be manually removed using “NOT”.

Tract #9: Uncinate fasciculus (UNC) (Fig.9)—The most posterior coronal slice in which the temporal lobe is separated from the frontal lobe is selected (Figs. 9a and 9c). The first ROI includes the entire temporal lobe and the second ROI includes the entire projections toward the frontal lobe.

Tract #10: The forceps major (Fmajor) (Fig.10)—It is difficult to devise a protocol that can reconstruct the entire corpus callosum reproducibly. In this paper, reconstruction of callosal connections to the occipital lobe *via* the splenium of corpus callosum (forceps major) and to the frontal lobe *via* the genu of corpus callosum (forceps minor) are devised.

For the 1st ROI for the forceps major (Fmajor), a para-sagittal plane at the level of cingulum is selected (Fig.7a) and the parieto-occipital sulcus is identified. A coronal plane is selected at the most posterior edge of this sulcus (coronal slice #3 in Fig. 7b) so that the coronal plane includes only the occipital lobe. The least-diffusion-weighted image may be used for better identification of the sulcus. The first and second ROIs delineate right and left occipital lobe as shown in Fig. 10a and 10c.

Tract #11: The frontal projection of the corpus callosum (the forceps minor) (Fminor) (Fig.11)—For the first ROI (Fig.11b), a coronal plane at the middle point between the anterior tip of frontal lobe and the anterior edge of the genu of the CC is selected using the mid-sagittal plane (Fig. 11a, between the two arrows). The first and second ROIs delineate the entire frontal lobe of each side as shown in Fig. 11b and 11d).

Reproducibility measurements

To quantify the reproducibility of the fiber tracking protocol, intra-rater and inter-rater reproducibility was assessed. The 11 protocols were repeated three times by an operator for the intra-rater reproducibility. For inter-rater reproducibility, the protocols were performed by three different operators using the same data set (called Subject #1 hereafter). Three raters participated within the same institute (S.W., P.D., and A.B at JHU). Protocols that passed the initial intra-institutional reproducibility tests (kappa values more than 0.6) were passed to 3 additional raters in 3 different institutes (A.C. at UNM, M.P. at UCSD, and M.P. at UCI), to measure inter-institutional reproducibility. The entire procedure was repeated using two more datasets from different subjects (Subject #2 and Subject #3) and the results were averaged. The 3 raters outside JHU didn't have prior experience in tractography and their training and instruction was accomplished independently using the written protocols. Therefore, the inter-institutional reproducibility could be considered as the poorest reproducibility condition; but the one that most accurately reflects the state of the field.

Statistical Analysis of spatial matching by κ

Spatial matching was examined for statistical analysis of reproducibility. The repeated tracking results were first converted to binary information of the same pixel dimension as the DTI data (256×256×50), in which pixels that were occupied by the tracts were assigned to a value of 1 and other non-occupied pixels were assigned to a value of 0. Two tracking results out of 3 trials were then superimposed, which yields to four different pixel categories; (1) pixels did not contain the tract in both trials (nn), (2) pixels that contained the tract in only one of the two trials (pn, np), and (3) pixels that contained the tracts in both trials (pp). Expectation values (Enn, Enp, Epn, and Epp) for each class were then calculated using the equations;

$$\text{Expected nn (Enn)} = (nn + np)(nn + pn) / N$$

$$\text{Expected np (Enp) or Epn} = (nn + np)(np + pp) / N \text{ or } (nn + pn)(pn + pp) / N$$

$$\text{Expected pp (Epp)} = (pn + pp)(np + pp) / N$$

Where $N=nn+np+pn+pp$ is the total number of pixels for the particular fiber. For the calculation, pixels with FA lower than the threshold ($FA > 0.2$) were not included. Then κ (kappa) was calculated by $\kappa = (\text{observed agreement} - \text{expected agreement}) / (100 - \text{expected agreement})$, where

$$\text{observed agreement} = (nn + pp) / N * 100$$

$$\text{expected agreement} = (Enn + Epp) / N * 100$$

This analysis was applied in a pair-wise manner; there are three combinations from the 3 trials. The κ values were determined for the three pair-wise combinations and an average κ was determined. According to criteria set by Landis and Koch (1977), the κ value of 0.11-0.2 is considered as “slight,” 0.21-0.4 as “fair”, 0.41-0.60 as “moderate”, 0.61-0.80 as “substantial”, and 0.81-1.0 as “almost perfect” agreement.

Study of tract-specific MR parameters in the normal subjects

The protocols with high reproducibility were applied to the normal population ($n = 10$) and coordinates of each of the white matter tracts were identified. These coordinates were superimposed on FA, and T_2 maps to measure these parameters in a tract-specific manner. The tract size was also calculated from the number of pixels occupied by the tracts.

These results were represented in two different ways. First, the all occupied pixels were pooled to determine the average parameter value for each tract. Second, the parameters were projected onto one of the anatomical axes (right-left, anterior-posterior, and superior-inferior) to observe changes in the parameters along the course of the trajectory. For inter-subject averaging, the tract length was normalized using anatomical landmarks along its course. For example, the CST was segmented into the midbrain, internal capsule, and corona radiata regions. The length of these separate regions from each subject was linearly adjusted before group averages were calculated.

Results

1: Intra and inter-rater reproducibility

Table 1 shows κ values of intra and inter-rater reproducibility (3 raters within an institute), inter-institutional reproducibility (3 raters, one each from three different institutes for Subject #1-1 and Subject #1-2 data), and average results for three different datasets (3 raters from three institutes for Subject #1, Subject #2, and Subject #3 data). For the 11 fibers with the established protocol, κ values are always higher than 0.6, and mostly higher than 0.7. As expected, inter-institutional reproducibility tends to be poorer than other reproducibility measures. Differences between the “AND” and “CUT” operations are unexpectedly small, with reproducibility of “CUT” operation often lower than that of “AND”. This is because the tract definition for the “CUT” operation is sensitive to the locations of ROIs, especially when two ROIs are close together (e.g. SLF). In these cases, a one-slice shift of the ROI location leads to addition or subtraction of a group of pixels within the slice and has a large effect on the κ value.

The intra and inter-rater reproducibility of FA measurements were also performed by superimposing each trial result on the FA map. The coefficients of variation ($CV = \text{standard deviation/average}$) for the intra and inter-rater variability are very small for all 11 tracts; the CV averaged over the 11 tracts are 0.8 and 1.5% for intra and inter-rater variability using AND operations and 1.1 and 1.0% using CUT operations.

2: Study of normal population

Averaged sizes of fibers (the number of pixels that contain the fibers multiplied by the nominal size of the pixel) for 10 healthy volunteers are shown in Fig. 12B. The reconstructed tracts tend to be larger in the left hemisphere and the differences are significant for the CGC, SLFt, ILF, and UNC, while no difference is found for CST and IFO. Interestingly all asymmetric tracts are related to the temporal lobe. Please note that these volumes were not normalized for the differences in cranial volumes. The asymmetry results, however, should not be influenced by the normalization. The asymmetry of FA is far smaller and significant difference is found only in CGC ($L > R$) and CGH ($R > L$) (Fig. 12A).

In Fig. 13, a correlation plot between T_2 and FA is shown for each tract. Most tracts are clustered in the T_2 range of 65 – 70 ms and FA range from 0.4 – 0.5. Exceptions are the corpus callosum, featured with a high anisotropy and the corticospinal tract, featured by high anisotropy and long T_2 although the differences were very small (0.02-0.04 in FA).

Data in Fig. 13 were calculated from all pixels that belong to a tract. The data can be projected to one of the anatomical axes. Fig. 14 shows results of several tracts with characteristic and reproducible variation of T_2 and FA along the course of their trajectories. For example, the CST has relatively high anisotropy in the internal capsule. Relatively long T_2 of the CST is apparent throughout its trajectories, although it has monotonic decrease toward the cortex. The IFO has an area with high FA at coronal slices near the anterior commissure and also in the sagittal stratum.

Discussion

Importance of protocol setup and measurement of reproducibility

Tractography is a unique tool that allows us to study white matter architecture three-dimensionally and non-invasively. It can quantitatively illustrate trajectories of various white matter tracts, which is useful not only for educational purposes, but also for advancing our understanding of abnormal brain anatomy. It has been also shown that tractography can be used for quantification studies. Once three-dimensional coordinates of a specific tract are identified, such coordinates can be superimposed on various MR parameter maps to quantify them. However, there are always two important questions that have been repeatedly asked. These are; 1) Precision: how reproducible the results are and 2) Accuracy: how real or valid the results are.

There are no general answers to these questions because they depend on the particular tracts of interest, imaging parameters, and tracking algorithms including choice of ROIs and thresholds for FA and angle). Furthermore, accuracy is not easy to measure because of the lack of a gold standard. On the other hand, with a given tract and reconstruction protocol, it is possible to measure its precision by repeating the measurement. Variability in reconstruction results stem from data acquisition (e.g. SNR, partial volume effect, patient motion, etc) and ROI placement for tracking. The former factors depend on imaging protocols and there are previous publications examining their impacts on the reconstruction results (Heiervang et al., 2006; Huang et al., 2004). The latter can be improved by refining ROI drawing protocols.

This precision (reproducibility) measurement could be one of the most important steps to establish the tractography as a useful research and clinical tool. Even if its accuracy is unknown, if tractography can detect reproducible difference in a specific tract between two groups, this would be an important information to understand disease status or even mechanism.

As for accuracy of tracking results, we reconstructed tracts that are well documented in previous anatomical studies using anatomical constraints (multiple ROIs) based on *a priori* knowledge. The macroscopic configuration of these reconstructed tracts are, thus, likely to reflect true fiber bundles. However, it is possible that some parts of the trajectory may contain inaccuracy due to partial volume effects, noise, and crossing fibers. It is also important to understand that the visualized pathways do not necessarily reflect brain connectivity because individual axons could be merging and blanching at any points along the bundle. In an extreme case, there may not be any axons that follow the entire path of the reconstructed trajectory.

Setup of reproducible tracking protocols involves two steps. First we need to know which tracts can be reconstructed reproducibly and second we have to refine protocols to reconstruct them. The former step depends on the size, curvature, and discreteness of the tracts. The second factor

involves iteration of protocol setup and reproducibility measurements. In our study, we initially identified about 20 known fiber tracts and could develop reproducible protocols for 11 of them. There are several important tracts that are not included in this paper such as fornix, stria terminalis, fimbria, anterior commissure, posterior commissure, body of the corpus callosum, superior thalamic radiation, and optic radiation. These fibers are not included because our current protocols have low reproducibility (κ less than 0.6) or we found large deviations among the healthy subjects. For those fibers with narrow, tubular trajectories (fornix, stria terminalis, anterior and posterior commissure), the reason for the poor reproducibility is most likely due to not enough spatial resolution with respect to the diameter. For the optic radiation, the curvature of tracts at the Meyer's loop is too steep to follow with 2.5 mm pixel resolution. The body of the corpus callosum and the superior thalamic radiation can readily be reconstructed but defining a specific protocol to reproducibly identify 2D slices for ROI drawing is challenging. We would like to emphasize that the reproducibility of tractography results shown in Table 1 are high because these protocols went through optimization procedures to improve the reproducibility and, as mentioned above, those with poor reproducibility were rejected through the process. Therefore, these figures should not be considered as an index of the reproducibility of tractography in general. It is also worth noting that given sufficient thought and technical refining as was done for these eleven tracts, many additional tracts could be reproducibly defined by a laboratory.

The protocols introduced in this paper are not the only reproducible protocols and their reproducibility may vary depending on imaging parameters. Therefore, they should be considered as a guideline. Using these protocols, we measured various MR parameters and size of each tract in normal subjects. The results of the normal variations should provide important information to perform power analysis and judge sensitivity to detect abnormalities.

Tract-specific quantification

As previously pointed out, (Kinoshita et al., 2005) we observed a significant amount of standard deviations in tract sizes. Because intra-rater and inter-rater variability or reproducibility of data acquisition are relatively small (Table 1), the large variability in the population study is likely to have originated from individual anatomical variability. Tractography integrates all events along the tract and thus is influenced by many factors. For example, if results indicate a smaller tract size in a subject, it could be due to the smaller size of the entire length of the tract or there could be one region where the tract is narrowed. Other reasons include perturbation of eigenvectors by crossing fibers and lower anisotropy. In other words, tractography is a method that is very sensitive to various anatomical variations along the trajectory, which leads to a large amount of variability among healthy subjects. It is, at this point, not clear if tractography-based measurement of tract volumes could be a sensitive method to detect anatomical abnormalities due to a pathological condition.

Despite of the large standard deviations of the reconstructed tract sizes, tract-specific T_2 and FA measurement results showed characteristic signatures of individual tracts with small standard deviations. The T_2 -FA correlation plot of could separate the corticospinal tracts and the corpus callosum from other tracts. The long T_2 of the CST agrees with previous studies (Yagishita et al., 1994). Once the normal signature of each tract is defined in this type of multi-modal plot, it could be a powerful tool to detect abnormalities in a tract-specific manner. In this respect, it could be an interesting effort to add more parameters such as MTR and T_1 . For more in-depth analysis of individual tracts, these MR parameters can be examined along its course as previously demonstrated. (Glenn et al., 2003; Pagani et al., 2005; Partridge et al., 2004; Stieltjes et al., 2001; Virta et al., 1999; Wilson et al., 2003; Xue et al., 1999) Some tracts have rather monotonous profile along their trajectories, while others, such as those shown in

Fig. 14, show characteristic and reproducible changes in MR parameters. This approach should provide detailed information to detect regional abnormalities.

The choice of “AND” and “CUT” operations depends on purpose of the study. The “AND” operation should be used when one is interested in characterizing the entire tract while “CUT” is suitable if one segment of a tract needs to be measured. Alternatively, the tract projection method (Fig. 16) could be used in which regions inside and outside the ROIs are spatially separated in the plot.

In Table 2, results of power analyses are tabulated based on our measurement of tract volumes and FA values of normal subjects. This could be used as a quick reference to estimate the number of subjects required to detect abnormalities. While this paper describes reproducibility of tractography procedures and variability in the normal population, we have not characterized image acquisition reproducibility, which is typically measured by scanning the same subject repeatedly. This is obviously a function of signal-to-noise ratio and thus directly related to imaging parameters such as scanning time and image resolution. A comprehensive study to evaluate the image acquisition reproducibility is currently being performed under the initiative of NIH/NCRF-funded Biomedical Informatics Research Network (BIRN, www.nbirn.net) (Farrell et al., 2006).

Asymmetry of tracking results

The finding of asymmetry in the size of fibers in the temporal lobe (SLF, UNC, ILF, and CG) is somewhat unexpected. At this point, we can't conclude which temporal regions account for this difference and whether the difference is caused by difference in anisotropy or actual size of these tracts. Although FA of each tract defined by tractography has low asymmetry, we can not immediately conclude that difference in FA is not the reason for the tract size asymmetry because tractography uses FA as a threshold; difference in FA may manifest as smaller tract size and not necessarily as low FA of the reconstructed tracts. We would like to see if our findings are reproducible in other sites. In this respect, it is noteworthy that the asymmetry of the cingulum (Gong et al., 2005a; Gong et al., 2005b) and the superior longitudinal fasciculus (Makris et al., 2005; Nucifora et al., 2005) has been reported in multiple publications. Park et al performed voxel-based analysis and reported higher FA of the cingulum bundle in the left hemisphere, but otherwise for the uncinate fasciculus. (Park et al., 2004)

Conclusions

Protocols to reconstruct 11 major white matter tracts are described. For these selected white matter tracts, high reproducibility was observed. The established protocols were then applied to a DTI database of adult normal subjects to study size, FA, and T_2 of individual white matter tracts, which revealed characteristic signatures of each tract. Asymmetry of tract size is observed for tracts projecting to the temporal lobe. Our protocol could be used as guidance for tractography-based tract-specific quantification.

Acknowledgements

Grants:

P. van Zijl NCRF P41 RR15241

S. Mori RO1AG20012, U24 RR021382

Reference

Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A. In vitro fiber tractography using DT-MRI data. *Magn. Reson. Med* 2000;44:625–632. [PubMed: 11025519]

- Catani M, Howard RJ, Pajevic S, Jones DK. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage* 2002;17:77–94. [PubMed: 12482069]
- Conturo TE, Lori NF, Cull TS, Akbudak E, Snyder AZ, Shimony JS, McKinstry RC, Burton H, Raichle ME. Tracking neuronal fiber pathways in the living human brain. *Proc. Natl. Acad. Sci. USA* 1999;96:10422–10427. [PubMed: 10468624]
- Farrell, J.; Landman, B.; Jones, C.; Smith, S.; Prince, J.; van Zijl, P.; Mori, S. Effects of diffusion weighting scheme and SNR on DTI-derived fractional anisotropy at 1.5T. *International Society of Magnetic Resonance in Medicine*; Seattle: 2006. p. 1075
- Glenn OA, Henry RG, Berman JI, Chang PC, Miller SP, Vigneron DB, Barkovich AJ. DTI-based three-dimensional tractography detects differences in the pyramidal tracts of infants and children with congenital hemiparesis. *J Magn Reson Imaging* 2003;18:641–648. [PubMed: 14635148]
- Gong G, Jiang T, Zhu C, Zang Y, He Y, Xie S, Xiao J. Side and handedness effects on the cingulum from diffusion tensor imaging. *Neuroreport* 2005a;16:1701–1705. [PubMed: 16189481]
- Gong G, Jiang T, Zhu C, Zang Y, Wang F, Xie S, Xiao J, Guo X. Asymmetry analysis of cingulum based on scale-invariant parameterization by diffusion tensor imaging. *Hum Brain Mapp* 2005b;24:92–98. [PubMed: 15455461]
- Heiervang E, Behrens TE, Mackay CE, Robson MD, Johansen-Berg H. Between session reproducibility and between subject variability of diffusion MR and tractography measures. *NeuroImage* 2006;33:867–877. [PubMed: 17000119]
- Huang H, Zhang J, van Zijl PC, Mori S. Analysis of noise effects on DTI-based tractography using the brute-force and multi-ROI approach. *Magn Reson Med* 2004;52:559–565. [PubMed: 15334575]
- Jellison BJ, Field AS, Medow J, Lazar M, Salamat MS, Alexander AL. Diffusion tensor imaging of cerebral white matter: a pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. *AJNR Am J Neuroradiol* 2004;25:356–369. [PubMed: 15037456]
- Jones DK, Horsfield MA, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magn. Reson. Med* 1999a;42:515–525. [PubMed: 10467296]
- Jones DK, Simmons A, Williams SC, Horsfield MA. Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn. Reson. Med* 1999b;42:37–41. [PubMed: 10398948]
- Kinoshita M, Yamada K, Hashimoto N, Kato A, Izumoto S, Baba T, Maruno M, Nishimura T, Yoshimine T. Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation. *NeuroImage* 2005;25:424–429. [PubMed: 15784421]
- Lazar M, Weinstein DM, Tsuruda JS, Hasan KM, Arfanakis K, Meyerand ME, Badie B, Rowley HA, Haughton V, Field A, Alexander AL. White matter tractography using diffusion tensor deflection. *Hum Brain Mapp* 2003;18:306–321. [PubMed: 12632468]
- Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness VS Jr, Pandya DN. Segmentation of subcomponents within the superior longitudinal fascicle in humans: a quantitative, in vivo, DT-MRI study. *Cereb Cortex* 2005;15:854–869. [PubMed: 15590909]
- Mori S, Crain BJ, Chacko VP, van Zijl PCM. Three dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annal. Neurol* 1999;45:265–269. [PubMed: 9989633]
- Mori S, Kaufmann WE, Davatzikos C, Stieltjes B, Amodei L, Fredericksen K, Pearlson GD, Melhem ER, Solaiyappan M, Raymond GV, Moser HW, van Zijl PCM. Imaging cortical association tracts in human brain. *Magn Reson Med* 2002;47:215–223. [PubMed: 11810663]
- Mori, S.; Wakana, S.; Nagae-Poetscher, LM.; van Zijl, PC. *MRI atlas of human white matter*. Elsevier; Amsterdam, The Netherlands: 2005.
- Nucifora PG, Verma R, Melhem ER, Gur RE, Gur RC. Leftward asymmetry in relative fiber density of the arcuate fasciculus. *Neuroreport* 2005;16:791–794. [PubMed: 15891571]
- Pagani E, Filippi M, Rocca MA, Horsfield MA. A method for obtaining tract-specific diffusion tensor MRI measurements in the presence of disease: application to patients with clinically isolated syndromes suggestive of multiple sclerosis. *NeuroImage* 2005;26:258–265. [PubMed: 15862226]
- Park HJ, Westin CF, Kubicki M, Maier SE, Niznikiewicz M, Baer A, Frumin M, Kikinis R, Jolesz FA, McCarley RW, Shenton ME. White matter hemisphere asymmetries in healthy subjects and in schizophrenia: a diffusion tensor MRI study. *NeuroImage* 2004;23:213–223. [PubMed: 15325368]

- Parker GJ, Stephan KE, Barker GJ, Rowe JB, MacManus DG, Wheeler-Kingshott CA, Ciccarelli O, Passingham RE, Spinks RL, Lemon RN, Turner R. Initial demonstration of in vivo tracing of axonal projections in the macaque brain and comparison with the human brain using diffusion tensor imaging and fast marching tractography. *NeuroImage* 2002;15:797–809. [PubMed: 11906221]
- Partridge SC, Mukherjee P, Henry RG, Miller SP, Berman JI, Jin H, Lu Y, Glenn OA, Ferriero DM, Barkovich AJ, Vigneron DB. Diffusion tensor imaging: serial quantitation of white matter tract maturity in premature newborns. *NeuroImage* 2004;22:1302–1314. [PubMed: 15219602]
- Pierpaoli C, Barnett A, Pajevic S, Chen R, Penix LR, Virta A, Basser P. Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *NeuroImage* 2001;13:1174–1185. [PubMed: 11352623]
- Poupon C, Clark CA, Frouin V, Regis J, Bloch L, Le Bihan D, Mangin JF. Regularization of diffusion-based direction maps for the tracking of brain white matter fascicules. *NeuroImage* 2000;12:184–195. [PubMed: 10913324]
- Pruessmann KP, Weiger M, Scheidegger MB, Boesiger P. SENSE: sensitivity encoding for fast MRI. *Magn Reson Med* 1999;42:952–962. [PubMed: 10542355]
- Stieltjes B, Kaufmann WE, van Zijl PCM, Fredericksen K, Pearlson GD, Mori S. Diffusion tensor imaging and axonal tracking in the human brainstem. *NeuroImage* 2001;14:723–735. [PubMed: 11506544]
- Virta A, Barnett A, Pierpaoli C. Visualizing and characterizing white matter fiber structure and architecture in the human pyramidal tract using diffusion tensor MRI. *Magn.Reson.Imaging* 1999;17:1121–1133. [PubMed: 10499674]
- Wakana S, Jiang H, Nagae-Poetscher LM, Van Zijl PC, Mori S. Fiber Tract-based Atlas of Human White Matter Anatomy. *Radiology* 2004;230:77–87. [PubMed: 14645885]
- Wiegell M, Larsson H, Wedeen V. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology* 2000;217:897–903. [PubMed: 11110960]
- Wilson M, Tench CR, Morgan PS, Blumhardt LD. Pyramidal tract mapping by diffusion tensor magnetic resonance imaging in multiple sclerosis: improving correlations with disability. *J Neurol Neurosurg Psychiatry* 2003;74:203–207. [PubMed: 12531950]
- Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC. Automated image registration: I. General methods and intrasubject, intramodality validation. *J Comput Assist Tomogr* 1998;22:139–152. [PubMed: 9448779]
- Xue R, van Zijl PCM, Crain BJ, Solaiyappan M, Mori S. In vivo three-dimensional reconstruction of rat brain axonal projections by diffusion tensor imaging. *Magn. Reson. Med* 1999;42:1123–1127. [PubMed: 10571934]
- Yagishita A, Nakano I, Oda M, Hirano A. Location of the corticospinal tract in the internal capsule at MR imaging. *Radiology* 1994;191:455–460. [PubMed: 8153321]

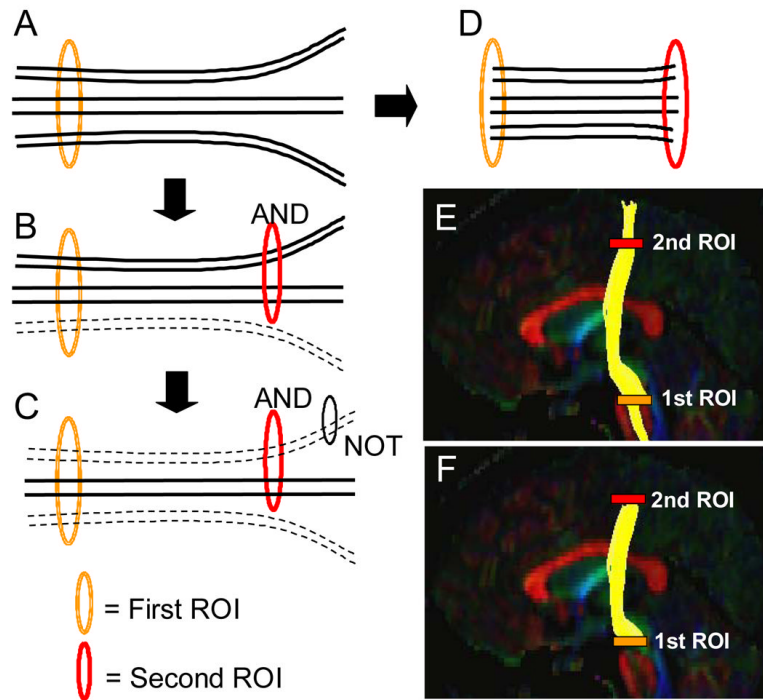


Figure 1.

A schematic diagram of three types of ROI operations; AND, NOT, and CUT. When the first ROI is drawn, all tracts that penetrate the ROI are retrieved (A). If the second ROI is applied as an "AND" operation, the fibers that penetrate both ROIs are retained (B). If a "NOT" operation is used, a subset of the fibers penetrating the NOT ROI is removed (C). The CUT operation is similar to the AND operation but only the tracking results between the two ROIs are retained (D). Results of "AND" and "CUT" operations are compared in (E: AND) and (F: CUT) using the corticospinal tract as an example.

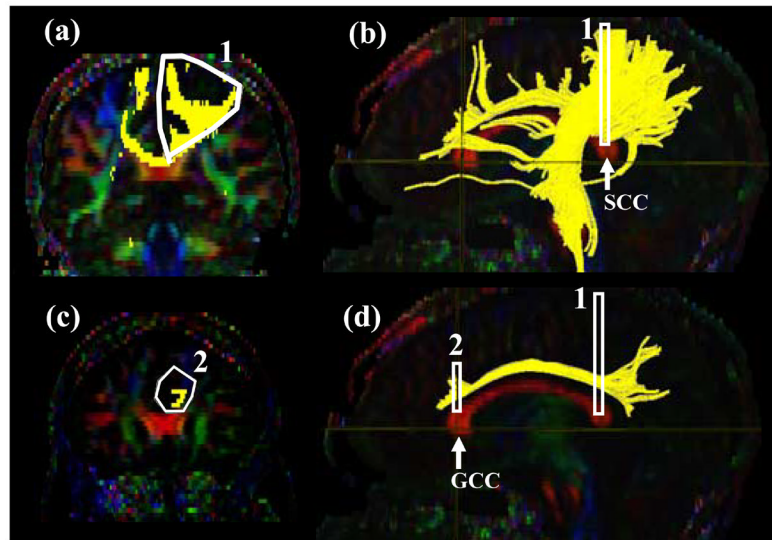


Figure 2. Locations of the ROIs for the cingulum in the cingulate gyrus part (CGC) on two coronal slices (a and c) and their locations in the mid-sagittal slice (b and d). The SCC and GCC stand for the splenium of corpus callosum and the genu of corpus callosum, respectively.

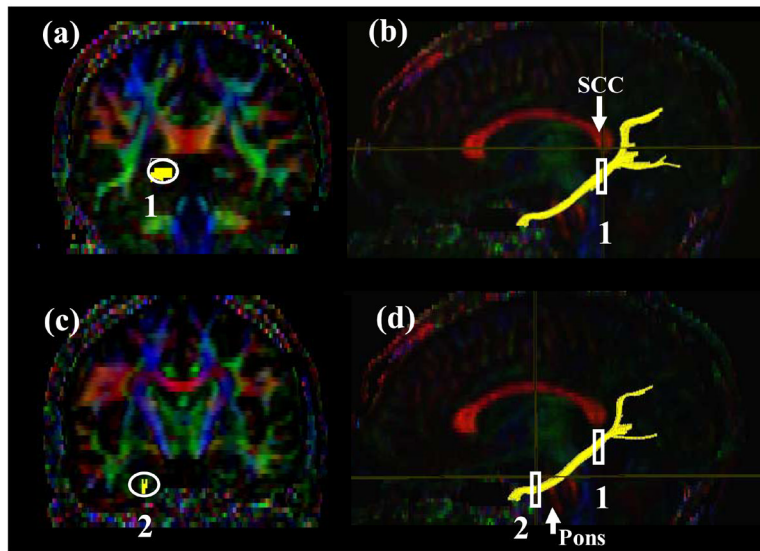


Figure3. Locations of the ROIs for the cingulum in the hippocampal part (CGH) on two coronal slices (a and c) and their locations in the mid-sagittal slice (b and d).

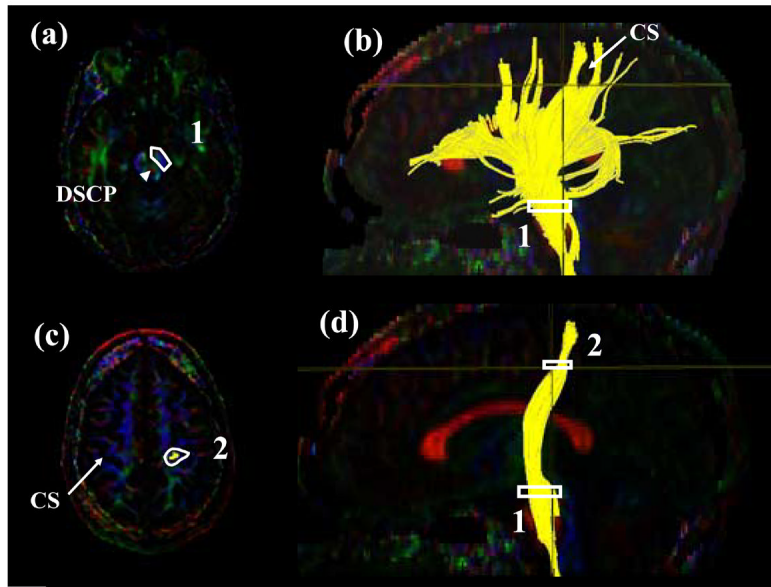


Figure.4.

Locations of the ROIs for corticospinal tract (CST) on two axial slices (a and c) and their locations in the mid-sagittal slice (b and d). The first ROI is drawn on the cerebral peduncle at the level of the decussation of the superior cerebellar peduncle (DSCP). From the tracking results, the central sulcus (CS) and the projection to the motor cortex are identified. Using the axial slice right after the bifurcation to the motor and sensory cortex, the CST is selected.

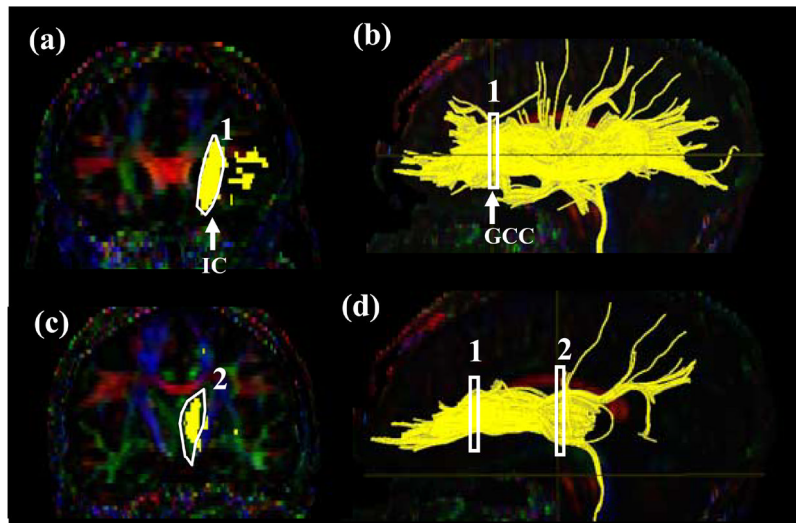


Figure 5. Locations of the ROIs for anterior thalamic radiation (ATR) on two coronal slices (a and c) and their locations in the mid-sagittal slice (b and d). The IC stands for the internal capsule.

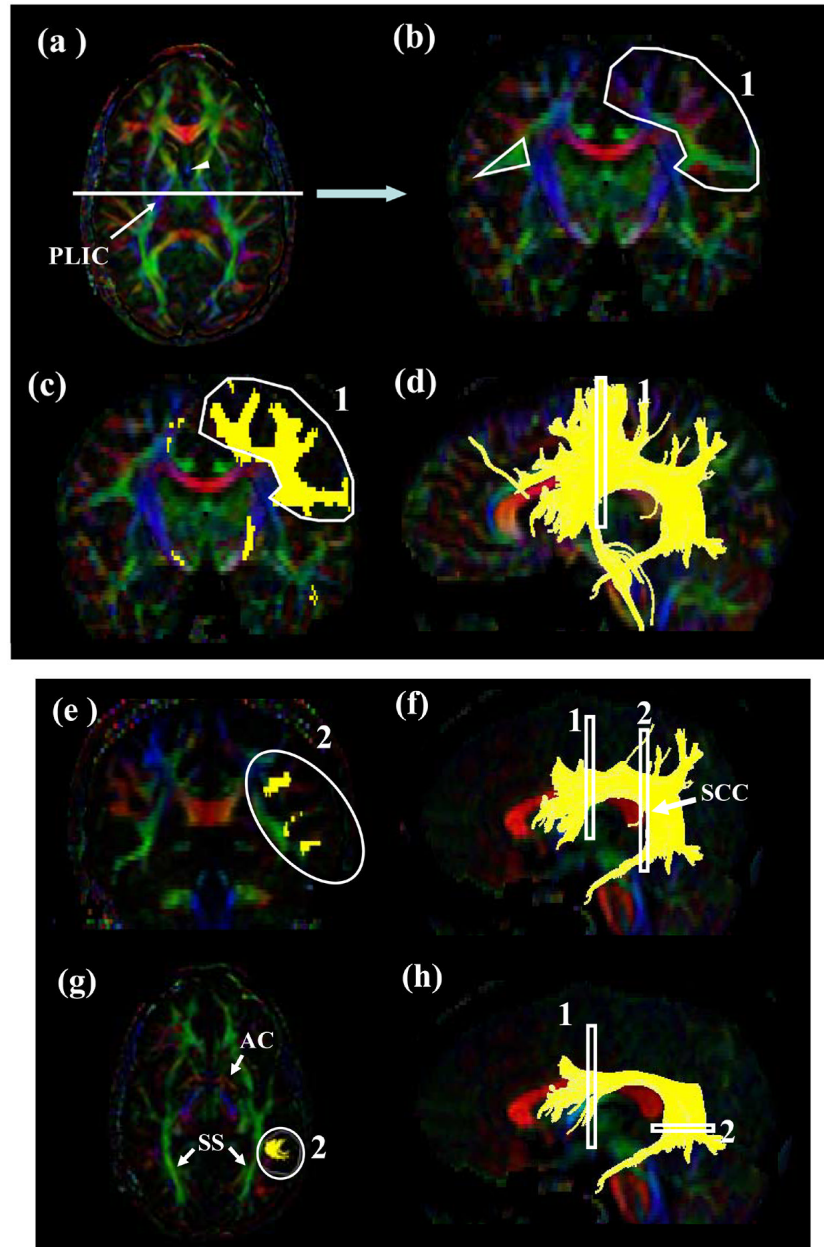


Figure 6. Locations of the ROIs for superior longitudinal fasciculus (SLF). At the middle of the posterior limb of the internal capsule (a, PLIC), a coronal slice is selected (b). The SLF can be identified as an intense triangle-shape green structure. The first ROI is shown in (c: coronal) and (d: sagittal). For the second ROI, a coronal slice (e) is selected at the splenium of corpus callosum (f). For the temporal component of the SLF (SLFt), the second ROI is drawn on the axial slice at the anterior commissure (AC) level (g and h). The SS stands for the sagittal stratum.

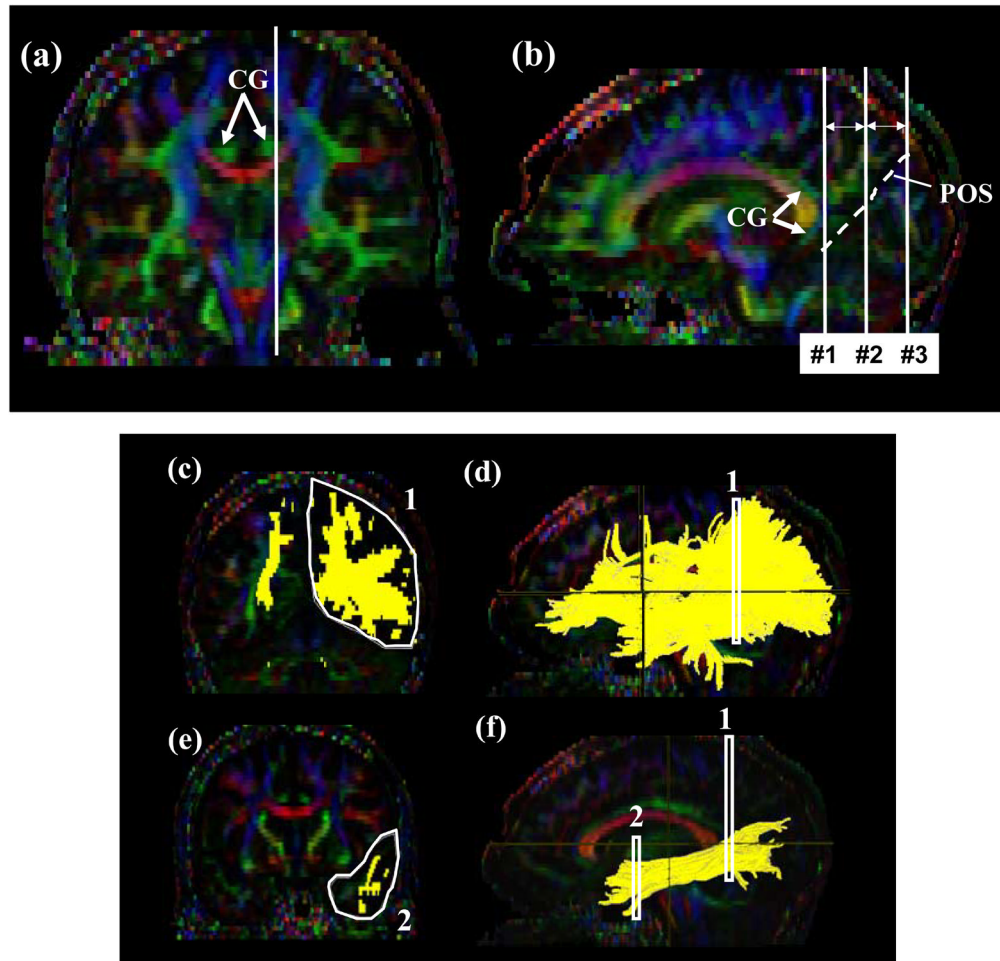


Figure 7.

Locations of the ROIs for the inferior longitudinal fasciculus (ILF). First a para-sagittal plane (b) is identified at the level of the cingulum (a, CG). The parieto-occipital sulcus (POS) is identified in the sagittal plane. The least-diffusion-weighted image could be used as a better anatomical guidance to identify the POS. A coronal slice (c and d) is selected at the posterior edge of the cingulum (slice #1). The second ROI defines trajectories toward the anterior pole of the temporal lobe (e and f).

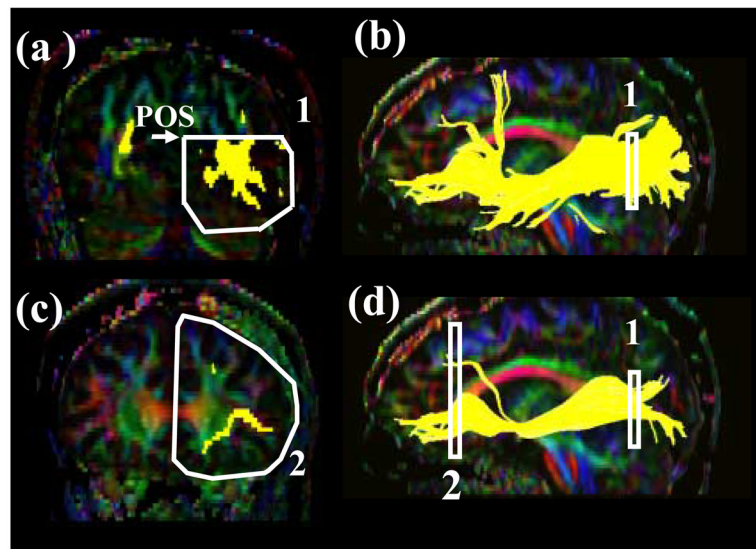


Figure 8. Locations of the ROIs for the inferior fronto-occipital fasciculus (IFO) on two coronal slices (a and c) and their locations in the mid-sagittal slice (b and d). For the coronal slice in (a), the #2 slice in Fig. 7b is used.

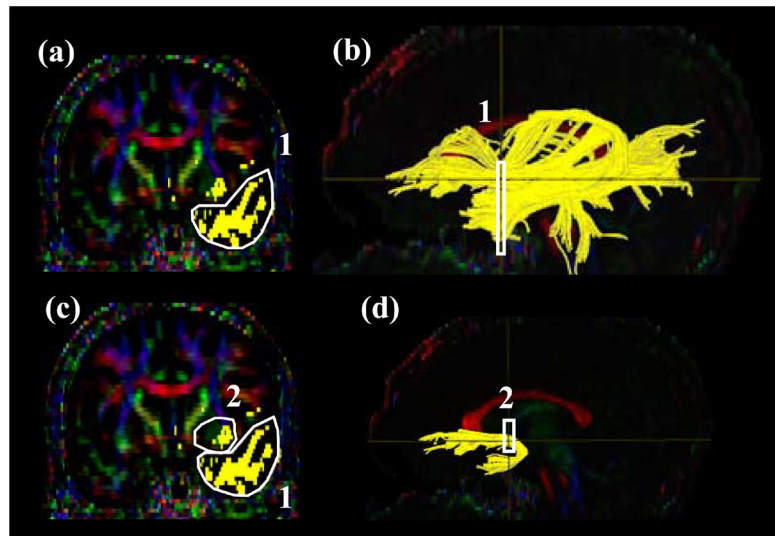


Figure 9.

Locations of the ROIs for the uncinatus fasciculus (UNC) on a coronal slice (a and c) and their locations in the mid-sagittal slice (b and d). The coronal slice (a and c) is the most posterior slice where the frontal and temporal lobe is separated. The least-diffusion-weighted image could be used for better anatomical guidance.

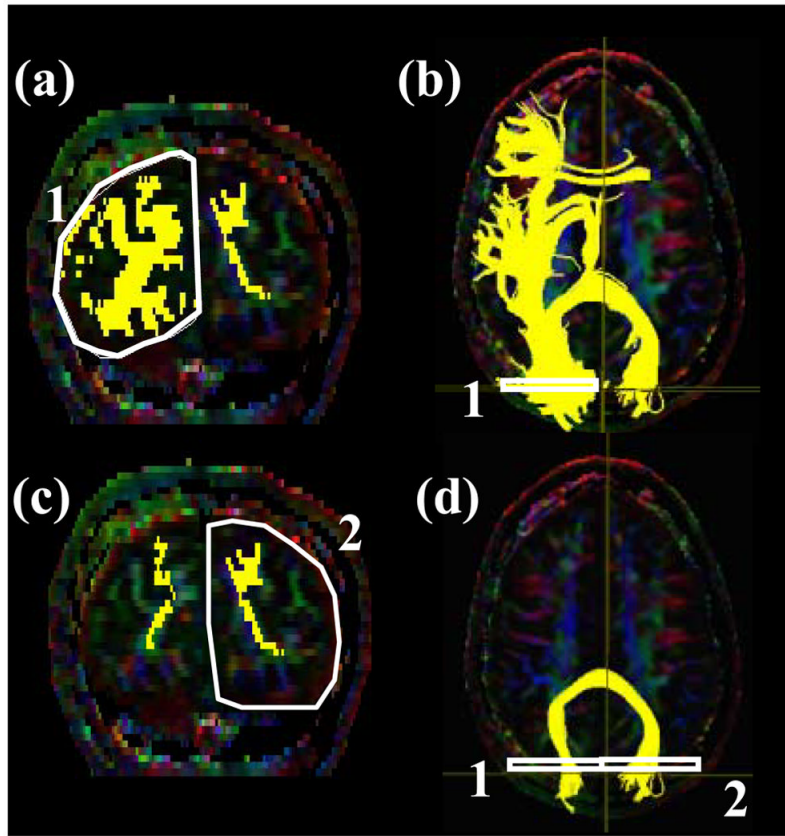


Figure 10. Locations of the ROIs for the occipital projection of the corpus callosum (forceps major, FMajor) on a coronal slice (a and c) and their locations in an axial slice (b and d). For the coronal slice in (a) and (b), the #3 slice in Fig. 7b (posterior edge of the POS) is used.

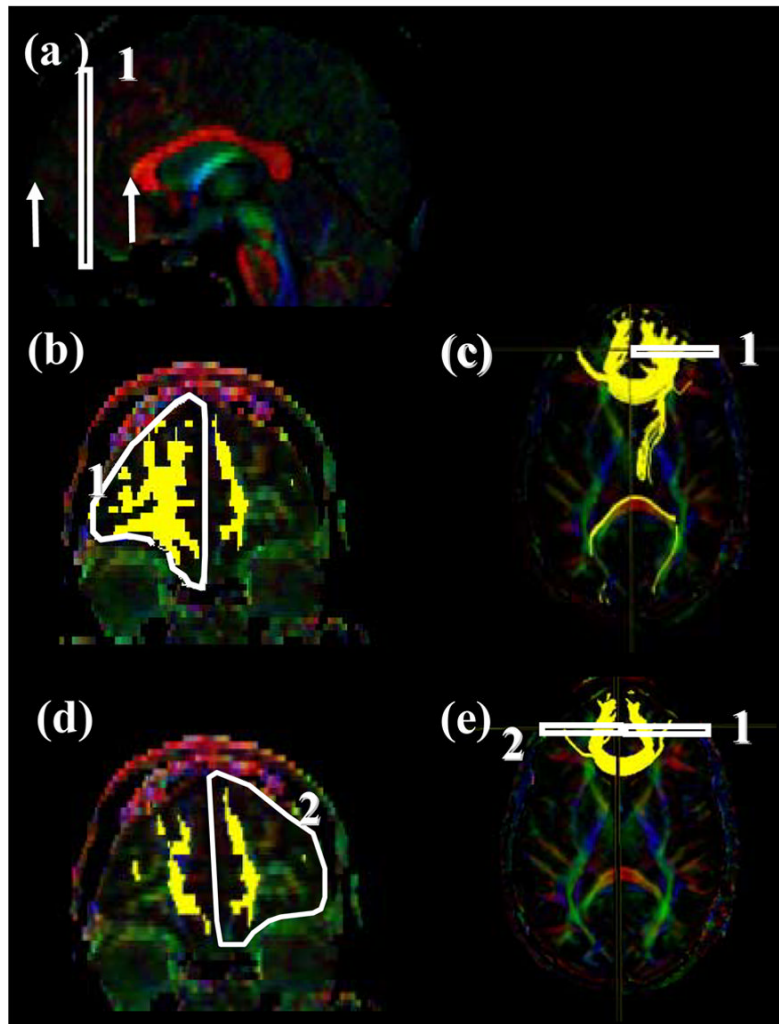


Figure 11. Locations of the ROIs for the frontal projection of the corpus callosum (forceps minor, FMinor) on a sagittal slice (a), a coronal slice (b and d) and their locations in an axial slice (c and e).

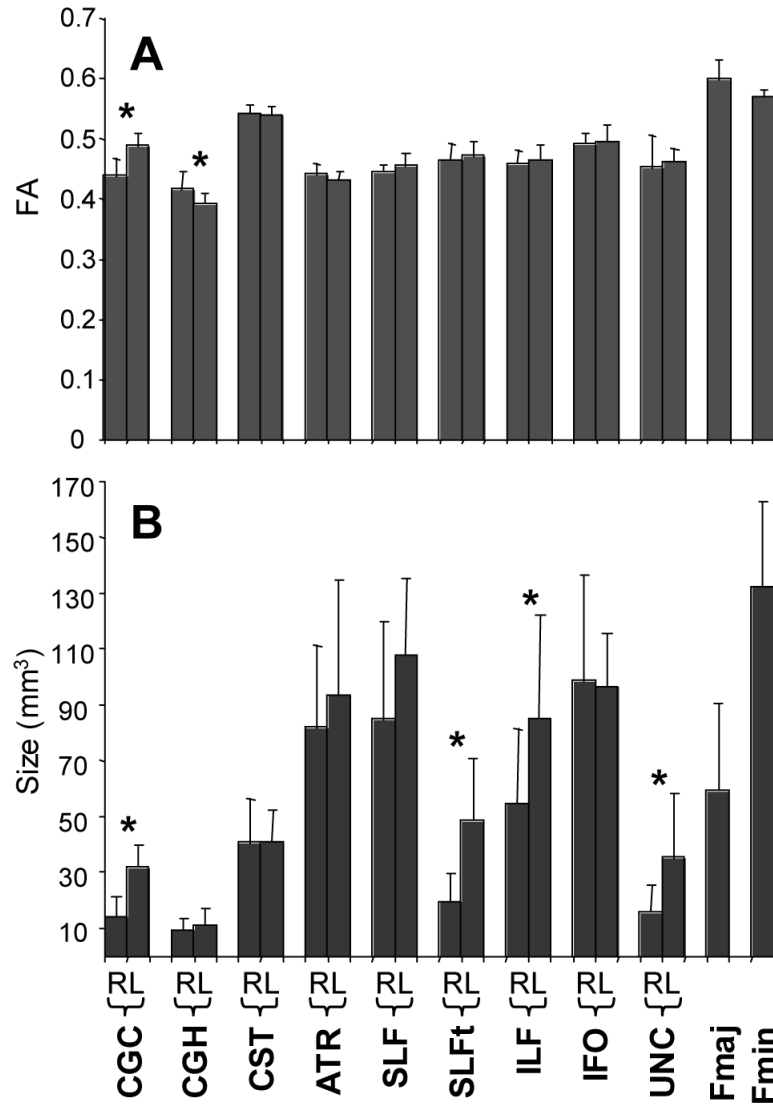


Figure 12.

FAs and sizes of white matter tracts of each reconstructed tract in 10 healthy volunteers.

Averages and standard deviations of the 10 subjects are shown (the standard deviation reflects the population variance, not the measurement reproducibility). Asterisks indicate statistically significant difference with $p < 0.05$. R and L indicate right and left hemisphere.

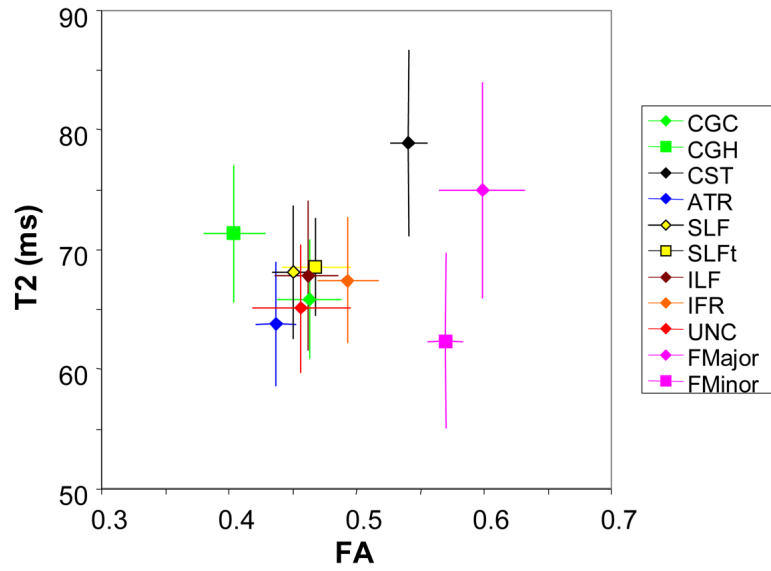


Figure 13. A correlation plot for FA and T₂ of each tract. Most of tracts are clustered in the FA range between 0.4-0.5 and a T₂ range between 60-75 ms, except for the CST, FMajor, and FMinor.

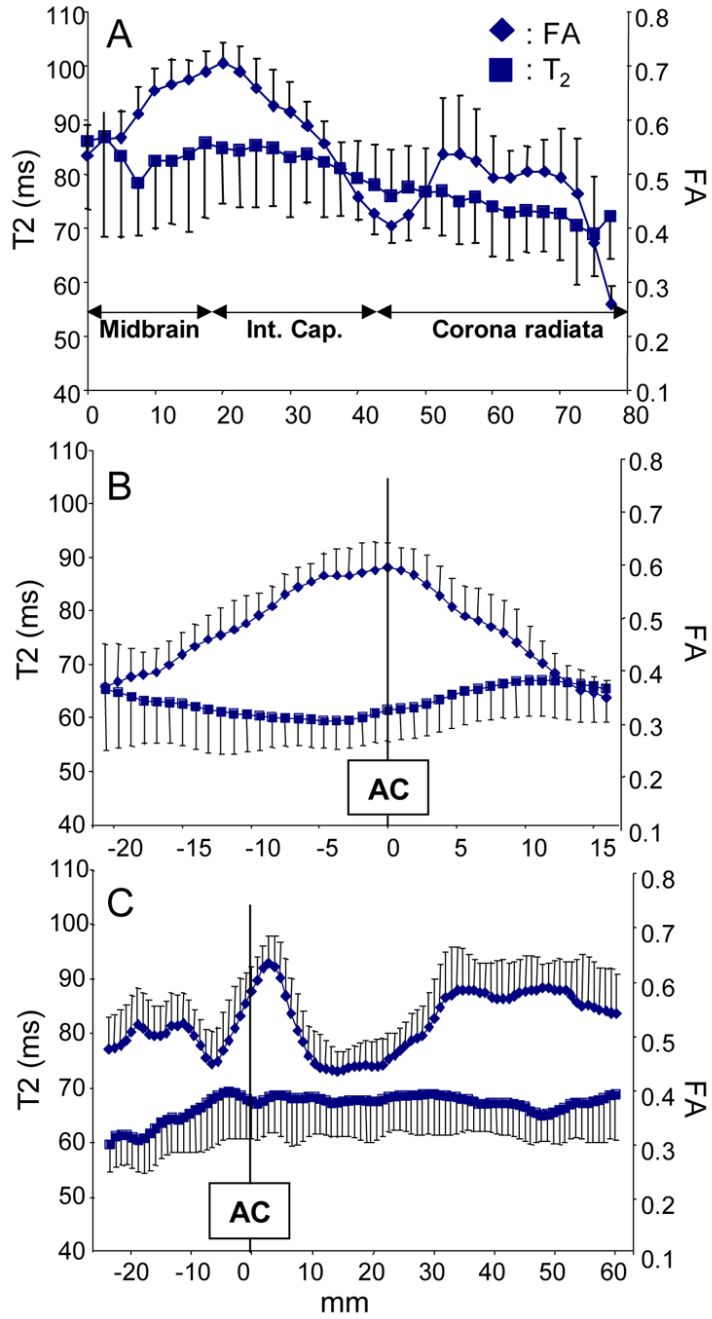


Figure 14. Tract-specific profiles of FA and T2 of the CST (a), the ATR (b), and IFO (c). Abbreviations are; Int. Cap.: internal capsule and AC: anterior commissure. Tract lengths from individual subjects were normalized based on more than 3 anatomical landmarks along their path.

Table 1

The k value of Intra- and Inter-rater Variability for the Eleven Major Tracts

	CGC	CGH	CST	ATR	SLF	SLFt	ILF	IFO	UNC	Fmaj	Fmin
Intra-rater	AND	0.994	0.992	0.94	0.902	0.944	0.974	0.985	0.996	0.976	0.983
	CUT	0.989	0.955	0.952	0.919	0.914	0.94	0.942	0.989	0.955	0.952
Inter-rater	AND	0.968	0.814	0.797	0.763	0.913	0.694	0.882	0.919	0.932	0.973
	CUT	0.940	0.911	0.793	0.797	0.888	0.654	0.893	0.912	0.966	0.975
Inter-Institute	AND	0.884	0.953	0.869	0.870	0.753	0.823	0.851	0.916	0.645	0.908
	CUT	0.879	0.900	0.787	0.787	0.695	0.732	0.842	0.810	0.600	0.860
3 subjects average	AND	0.873	0.831	0.889	0.831	0.745	0.782	0.856	0.758	0.735	0.911
	CUT	0.788	0.818	0.854	0.702	0.741	0.754	0.815	0.800	0.690	0.873

0.81-1.0 ; almost perfect agreement, 0.61-0.80 ;substantial, 0.41-0.60 ; moderate, 0.21-0.4 ; fair , 0.01-0.2 ; slight

Table 2
Power analysis to detect 10, 20, and 40 % changes with $p = 0.05$ and the power of 0.8.

	CGC	CGH	CST	ATR	SLF	SLFt	ILF	IFO	UNC	Fmaj	Fmin
10%	Volume	518*	337	360	525	342	813	298	1354	872	172
	FA	14	12	6	8	11	11	12	29	12	4
20%	Volume	126	87	93	130	96	173	77	341	220	46
	FA	6	6	4	4	4	4	4	10	4	4
40%	Volume	34	55	27	36	27	47	27	85	58	12
	FA	4	4	4	4	4	4	4	6	4	4

Numbers may vary depending on SNR and quality of datasets

* The number of subjects required to detect the specified change (averaged values for the right and the left hemisphere). The reported numbers were obtained from DTI datasets with 12 min scanning time.