Published in final edited form as:

Proc SPIE. 2013 March 13; 8669: . doi:10.1117/12.2007130.

Effects of DTI spatial normalization on white matter tract reconstructions

Nagesh Adluru^a, Hui Zhang^b, Do P. M. Tromp^a, and Andrew L. Alexander^a

^aWaisman Center, University of Wisconsin-Madison, USA

^bDepartment of Computer Science, University College London, UK

Abstract

Major white matter (WM) pathways in the brain can be reconstructed *in vivo* using tractography on diffusion tensor imaging (DTI) data. Performing tractography using the native DTI data is often considered to produce more faithful results than performing it using the spatially normalized DTI obtained using highly non-linear transformations. However, tractography in the normalized DTI is playing an increasingly important role in population analyses of the WM. In particular, the emerging tract specific analyses (TSA) can benefit from tractography in the normalized DTI for statistical parametric mapping in specific WM pathways. It is well known that the preservation of tensor orientations at the individual voxel level is enforced in tensor based registrations. However small reorientation errors at individual voxel level can accumulate and could potentially affect the tractography results adversely. To our knowledge, there has been no study investigating the effects of normalization on consistency of tractography that demands *non-local* preservation of tensor orientations which is not explicitly enforced in typical DTI spatial normalization routines. This study aims to evaluate and compare tract reconstructions obtained using normalized DTI against those obtained using native DTI. Although tractography results have been used to measure and influence the quality of spatial normalization, the presented study addresses a distinct question: whether non-linear spatial normalization preserves even long-range anatomical connections obtained using tractography for accurate reconstructions of pathways. Our results demonstrate that spatial normalization of DTI data does preserve tract reconstructions of major WM pathways and does not alter the variance (individual differences) of their macro and microstructural properties. This suggests one can extract quantitative and shape properties efficiently from the tractography data in the normalized DTI for performing population statistics on major WM pathways.

1. INTRODUCTION

Despite many recent developments in higher angular resolution modeling of diffusion weighted magnetic resonance imaging (DWI), diffusion tensor imaging (DTI) is still one of the most widely used models in neuroimaging research studies of white matter (WM) (1). Fiber tractography is a powerful algorithmic framework that can be used to reconstruct WM pathways in the brain using DTI data (2; 3). A specific quality of tractography is that it allows for extracting qualitative and quantitative information on the shape and structure of the WM pathways. This information may be utilized surgical planning and may also aid clinical and research studies (4). One can also use tractography data for obtaining various superresolution contrasts of the WM under the tract density imaging framework (5). Despite the challenge of quantitation using tractography (6), there are three popular and strong clinical applications (or population studies), namely [1] delineating specific pathway or region of interest in the white matter to test region specific hypotheses using various DTI

metrics such as fractional anisotropy (FA), mean diffusivity (MD) in those regions and [2] studying various shape characteristics of the pathways for applications such as classification (7), medial axes based tract specific statistical mapping (8) and [3] for structural network modeling (9; 10).

In all such applications a key step is establishing correspondences across the subjects in a study. Spatial normalization is a widely used computationally efficient framework of image registration for achieving this. A key step in spatially normalizing data is applying spatial transformations to the data, the other being *estimating* the transformations themselves. The primary contrast between spatially transforming DTI and other modalities of MRI such as T_1 weighted images is that the DTI data need to be reoriented to produce anatomically consistent results. (11) presented different strategies of reorientation for different types of transformation such as rigid, affine and diffeomorphic. For rigid transformations the tensors can be reoriented perfectly by just applying the rotation of the rigid transformation to the tensor. For affine and diffeomorphic transformations (which can be approximated as piecewise affine), (11) proposed two strategies namely finite strain (FS) which is spatially invariant and preservation of principle direction (PPD) which is spatially dependent based on the first eigen vector in the voxel. Since PPD explicitly enforces preservation of the first eigen vector it produces more accurate reorientation compared to FS (11). However when the transformations are small FS is nearly as accurate and permits computational efficiency (12). In fact it has been empirically shown in (13) that FS reorientation can be used when estimating spatial transformations and PPD reorientation can be used when applying the transformations to the tensors and this can result in accurate shape averaging for increased statistical sensitivity in voxel based analyses (VBA) to patterns in microstructural properties such as fractional anisotropy (FA).

However the effect of such tensor reorientation on tractography has not been studied. Although reorientation using PPD (or FS) can result in adequate accuracy for VBA (11; 13), since the commonly used algorithmic framework for tractography is sensitive to error accumulation (14), minor local errors can accumulate and may potentially cause undesirable effects on tractography based reconstructions of the pathways. Hence, performing tractography on the acquired diffusion tensor imaging (DTI) data, in the native space, is often considered to be more faithful to the white matter (WM) anatomy than performing it in the normalized space using non-linearly transformed data. The latter is considered less favorably because it requires spatial transformation of DTI data that involves the interpolation and reorientation of diffusion tensors, the effects of which on tractography has yet to be studied. If the topology and tensor orientation is properly maintained by using appropriately designed metrics (12), it should be possible to perform tractography on spatially normalized data with faithful results to those obtained in the native space. To our best knowledge, there has been no systematic study presenting data to confirm this hypothesis and presenting the effects of normalization, with local tensor reorientation, on the anatomical consistency of tractography based ROI extraction. There are several important reasons why testing the above hypothesis is important. To list a few, [1] tractography in the normalized space is playing an increasingly important role in population analyses of white matter. In particular, the emerging tract specific analyses such as (8, 15), which parameterize the medial representations of the tracts for population statistics and (16) can benefit from performing tractography in the normalized space. [2] In applications such as tractography-guided spatial statistics (TGIS) performing tractography directly in the normalized space can be computationally more efficient compared to a few iterations of warping the seed ROIs and tractography (17). [3] In frameworks such as tract based morphometry (18) one can avoid the tract clustering step since tracking in normalized space already produces corresponding tracts across the subjects. [4] Tractography reproducibility studies such as (19) use normalized space ROI tracing and perform tracking in the normalized space for

consistency. Our data will strengthen the results of such findings. [5] Finally even applications which prefer tracking in the native space typically identify waypoint and exclusion ROIs on an average in normalized space and then inverse warp those into the native space. Since identifying such ROIs following protocols described in (2, 3) can benefit tremendously when using tractography data, it is important to be able to know that performing tractography in the normalized space produces faithful results.

We would like to note that tractography results have been used to measure and influence the quality of DTI spatial normalization (12; 20; 21; 22) and that tractography reproducibility studies such as (19, 23) have been performed, but the work presented in this study addresses a distinct question: whether non-linear spatial transformations of DTI with non-exact tensor reorientation preserve even long-range anatomical connections when performing tractography in the normalized space. Hence, the objective of this study is to evaluate and compare tractography results obtained in normalized space versus those in native space. The two key hypotheses tested in this study are: [1] Quantitative summary measures, such as average fractional anisotropy (FA), mean diffusivity (MD) and total volume (VOL), of tracts obtained in both native and normalized spaces are consistent. [2] Tracts generated in both normalized and native space have a high degree of macrostructural anatomical shape consistency. We compare and evaluate the consistency of tractography based ROI extractions in normalized space and in native space using eight different commonly conceived DTI processing streams. Preliminary versions of the study discussing limited number of findings were presented in (24, 25).

2. MATERIALS AND METHODS

2.1 White Matter Tract Reconstructions

MR Acquisition and Basic Pre-processing—The diffusion weighted images (DWI) for four healthy adult subjects were acquired using a 3.0-Tesla GE Discovery MR750 MRI scanner (GE Healthcare, Waukesha, WI), the product 8-channel array head coil and a diffusion-weighted, spin-echo, echo planar imaging sequence. DTI measurements were performed using 48 non-collinear diffusion encoding directions with diffusion weighting factor of $b=1000$ s/mm² in addition to 8 $b=0$ images. Images were obtained with 2mm isotropic spatial resolution (field-of-view=256mm and 128×128 acquisition matrix) and were interpolated to 1mm×1mm in-plane using zero-filled interpolation on the scanner. Other parameters were TR=8s; TE=66.2ms; and echo spacing=684μs and flip angle = 60°. The DTI protocol was repeated four times, but not averaged on the scanner. Field maps were also acquired using a 2D spoiled gradient echo (SPGR) sequence at two echo times (TE1=7ms; TE2=10ms) to estimate the magnetic field strength across the brain and correct for geometric distortions in the DTI data. All the data were corrected for the eddy current related distortion and head motion using FSL software package (26). The distortions from field in-homogeneities were corrected using corresponding field maps. The brain tissue was extracted using the brain extraction tool (BET), part of the FSL software package. The tensor volumes are estimated using non-linear estimation using CAMINO (27), for all the 16 individual repeat scans as well as by combining four repeats as shown in Fig. 1(a).

DTI Spatial Normalization—The tensor volumes estimated from the individual repeats as well as from the combined scans were spatially normalized to a common DTI template which was obtained from scans (with equal acquisition parameters) of a different set of 16 healthy subjects from a similar age group. This mimics a standard setting of registering a population to a standardized template. This procedure is summarized in Fig. 1(b). DTI-TK (12) was employed for the purpose, as it has been shown to be an effective registration method for DTI data (28).

DTI Processing Streams—Extracting regions of interest (ROI) on diffusion tensor imaging (DTI) data is useful in testing specific univariate hypotheses as well as in multivariate pattern analyses. Although pre-defined white matter atlas ROIs such as Eve provide useful delineation of white matter regions, some ROIs such as uncinate fasciculus (a very important pre-frontal pathway in emotion regulation), can have limited coverage of the tract. Tractography on the other hand allows one to extract white matter ROIs with full coverage and on study specific data using protocols such as (3). But in a typical population analysis setting one is faced with several choices of DTI processing streams for extracting the ROI measurements. One example being whether to perform tractography in the native or spatially normalized space. In this paper we present data showing the consistency of tractography based ROI extractions using eight different commonly conceived DTI processing approaches. Hence we generate the following sets of processed data per subject as shown in Fig. 1 which summarizes the most commonly conceived ways of processing the DTI data. The **top** panel reflects the data generated by combining the repeated acquisition while the **bottom** reflects the same for individual repeats. The **left** of the dotted black line shows the data in native space and the **right** side shows the corresponding data in the normalized space. The dotted red arrows represent the application of spatial transformations to tensor volumes as well as the tractography data. The color-coding used is as follows: [1] Dark-orange indicates tensor volumes in native space and light orange indicates warped tensor volumes. [2] Dark-green indicates tractography performed on tensor volumes. Light-green reflects the warped tractography data. [3] Dark-magenta reflects the quantitative summaries of properties such as fractional anisotropy (FA), mean diffusivity (MD) or volume (VOL) obtained in native space while the light-magenta is used to indicate those in the normalized space. The quantitative summaries ($\bar{\mu}(\mathcal{T})$) are obtained by rasterizing the tractography data into binary masks and then applying those masks to the corresponding metric volumes computed in respective spaces. The various symbols also summarized in the Table 1 below.

Tractography—For each individual repeat (16 in total) and each combined scan (4 in total), we reconstructed 4 major WM pathways using tractography performed with both the original (s) and spatially normalized tensor volumes ($\mathcal{R}(s)$). The tractography was performed on the whole brain using FACT algorithm implemented in CAMINO (27). The seed file for tractography was the white matter mask defined as $FA > 0.2$ and the stopping criteria were $FA < 0.2$ and a maximum curvature threshold of 60° . The pathways examined were corpus callosum (CC), bilateral cingulum bundle (CB), bilateral inferior fronto-occipital fasciculus (IFO) and bilateral uncinate fasciculus (UF). These pathways were extracted by systematically following protocols described in (2, 3). Since the waypoint and exclusion filter regions (FRs) are difficult to delineate accurately using FA maps alone, we use the tractography data to iteratively refine the FRs. We use Trackvis software (29) for this purpose. Each pathway was extracted by applying the FRs to whole brain tractography i.e.,

$$\mathcal{T}_{\text{pathway}}(s) = \mathbb{F}(\mathcal{T}_{\text{FB}}(s), \text{FR}_{\text{pathway}}(s)), \quad (1)$$

where $\mathbb{F}$ denotes filtering of tracts using a set of waypoint and exclusion FRs needed for the specific pathway and $\mathcal{T}_{\text{FB}}(s)$ denotes full-brain (FB) tractography obtained on the tensor volume s . For all the pathways rather than tracing the required waypoint and exclusion FRs in the individual subject, they are first identified on the population average or atlas ($\mathcal{A}$) obtained by averaging the spatially normalized individual repeat data,

$$\mathcal{A} = \frac{\sum_{i=1}^4 \sum_{j=1}^4 F_i^j(s_i^j)}{16}. \quad (2)$$

These filter regions are then either inverse warped into the native space or applied directly in the normalized space. For example, the uncinate fasciculus tracts in the $\bar{s}_i$ and $F_i(\bar{s}_i)$ are obtained respectively as

$$\mathcal{T}_{UF}(\bar{s}_i) \equiv \mathbb{F}\left(\mathcal{T}_{FB}(\bar{s}_i), F_i^{-1}(\text{FR}_{UF}(\mathcal{A}))\right), \quad (3)$$

$$\mathcal{T}_{UF}(F_i(\bar{s}_i)) \equiv \mathbb{F}\left(\mathcal{T}_{FB}(F_i(\bar{s}_i)), \text{FR}_{UF}(\mathcal{A})\right). \quad (4)$$

This minimizes the amount of manually introduced inconsistencies and variability in the comparisons. Below we describe the tracts and their extraction briefly.

- [1] **Corpus Callosum:** One of the largest white matter pathways connecting the two cerebral hemispheres of the brain. Although these cortical connections are divided into three parts namely genu (the anterior projections), body (the middle part) and the splenium (the posterior projections), they can be reconstructed by placing a single waypoint regions in the midsagittal plane. CC also has extensions to the temporal lobe called tapetum which were constructed by creating additional waypoint regions in the temporal lobe and some exclusion regions in the inferior part midsagittal plane as shown in Fig. 2(a).
- [2] **Cingulum Bundle:** It belongs to limbic system fibers and has projections from temporal lobe and the nearby cingulate gyri both on the left and right sides of the brain. In this paper we reconstruct only the projections to the cingulate gyri using a cigar shaped seed/waypoint FR on the FA map using coronal view as shown in Fig. 2(b).
- [3] **Inferior Fronto-Occipital Fasciculus:** This tract shown in Fig. 2(c) belongs to the family of association fibers that occupies the ventral area of the external capsule connecting the orbitofrontal cortex and the occipital lobe. In the more posterior sections it starts to merge with the inferior longitudinal fasciculus. The two waypoint regions are identified on the atlas in the occipital lobe and the external capsule.
- [4] **Uncinate Fasciculus:** It belongs to the family of association fibers (also considered to belong to the limbic system) with significant projects to the fronto-orbital cortical areas from the temporal lobe. These are short range U shaped fibers and are one of the major amygdalo-prefrontal pathways. Delineation of the UF required the definition of filter regions by identifying the most posterior coronal slice that showed clear separation of the frontal and temporal lobes bilaterally. As depicted in Fig. 2(d) the FRs were manually drawn around the frontal lobes and around the temporal lobe on the identified coronal view of FA on the atlas ($\mathcal{A}$). These are identified on both left and right hemispheres of the brain. Then, the Boolean AND term was used to only include fibers that crossed through the frontal and temporal FRs for each hemisphere separately.

2.2 Consistencies of Tract Reconstructions

Since the two most common analyses using specific white matter pathways are either region-of-interest (ROI) analyses or shape based tract specific analyses (TSA), we perform two sets of quantitative evaluations that measure consistencies of quantitative summaries and shapes of the tracts when performing spatial transformations. The quantitative summary consistency results are most relevant for ROI analyses while the shape comparisons matter most for TSA and other feature extraction based analyses. For both sets of comparisons we use the data obtained using combined repeats ($\bar{s}$) to provide the baseline (gold-standard) reconstructions. The two sets of evaluations require different combinations of comparisons of the tractography data.

Consistency of Quantitative Summaries—In ROI analyses the average of the diffusion tensor properties (such as FA, Tr) of the tracts are used for statistical testing. Some times even the average volume (Vol) of the tracts are also used. Hence for these analyses we compare seven different methods of obtaining the averages against the gold-standard method as shown in Fig. 3(a). The comparisons 1 through 4 are using the average measurements obtained in the normalized space while those from 5 to 7 use the native space measurements. Since we do not have “ground-truth” tract reconstructions and compare the seven different methods against the baseline method ($\bar{\mu}(\mathcal{A})$), we report Bland-Altman (BA) plots (30) and concordance correlation coefficients, rather than correlation plots. BA plots are used to measure agreements between two methods both of which can have sources of uncertainty. The BA plots show scatter plots of the individual subject bias on the ordinate and the individual subject average of the two methods on the abscissa,

$$BA_i(x, y) = \left(\frac{m_1(i) + m_2(i)}{2}, m_1(i) - m_2(i) \right), \quad (5)$$

where $m_1(i)$ is the average FA/MD or log(VOL) for subject i obtained using one of the seven different methods and $m_2(i)$ is that obtained using the tractography obtained by the baseline method. These plots can be used to reflect the agreement between measurements obtained by different methods. If there is no systematic variance in the bias according to the average and if the individual subject biases in the metrics cluster around the total bias then there is high degree of agreement between two different methods. To provide a single number reflecting the degree of agreement we also report the concordance coefficient (ρ_c) which is defined as,

$$\rho_c = \frac{2\rho\sigma_{m_1}\sigma_{m_2}}{\sigma_{m_1}^2 + \sigma_{m_2}^2 + (\mu_{m_1} - \mu_{m_2})^2}, \quad (6)$$

where ρ is the Pearson correlation coefficient and μ_{m_1} and σ_{m_1} are the mean and standard deviations of the average microstructural properties obtained using one of the seven different methods. μ_{m_2} and σ_{m_2} reflect those obtained using the baseline method.

Consistencies of Geometric Shape—For the emerging white matter analyses such as tract specific analyses (TSA) the geometric shape features of the tracts become the interesting source of variance. Hence we measure the consistency of the shapes of the structures using dice similarity coefficients and scores. These evaluations shown in Fig. 3(b), reflect the macrostructural consistency of tractography results. Although there are myriad number of ways for capturing shape consistencies, dice similarity coefficients and scores are widely used in comparing shapes in constrained settings such as brain image analysis. The dice similarity coefficient between the pairs of masks obtained from the two sets of tractography, $\mathcal{T}_1, \mathcal{T}_2$ are computed as

$$DC(\mathcal{T}_1, \mathcal{T}_2) = \frac{2M(\mathcal{T}_1) \cap M(\mathcal{T}_2)}{|M(\mathcal{T}_1)| + |M(\mathcal{T}_2)|}, \quad (7)$$

where $M(\mathcal{T}_1)$ and $M(\mathcal{T}_2)$ are the binary masks obtained by rasterizing the tractography sets $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. The score, which is defined based on agreement between expected and observed data (31) is computed as

$$\kappa(\mathcal{T}_1, \mathcal{T}_2) = \frac{\text{obs} - \text{exp}}{1 - \text{exp}}, \quad \text{where} \quad (8)$$

$$\text{obs} = \text{nn} + \text{pp}, \quad \text{exp} = \text{Enn} + \text{Epp}, \quad (9)$$

$$\text{nn} = |M(\mathcal{T}_1) = 0 \ \& \ M(\mathcal{T}_2) = 0|, \quad \text{pp} = |M(\mathcal{T}_1) = 1 \ \& \ M(\mathcal{T}_2) = 1|, \quad (10)$$

$$\text{np} = |M(\mathcal{T}_1) = 0 \ \& \ M(\mathcal{T}_2) = 1|, \quad \text{pn} = |M(\mathcal{T}_1) = 1 \ \& \ M(\mathcal{T}_2) = 0|, \quad (11)$$

$$\text{Enn} = \frac{(\text{nn} + \text{pn})(\text{nn} + \text{np})}{N}, \quad \text{Epp} = \frac{(\text{pp} + \text{pn})(\text{pp} + \text{np})}{N}, \quad (12)$$

$$N = \text{nn} + \text{np} + \text{pn} + \text{pp}. \quad (13)$$

Dice coefficient >0.7 can be considered to imply a good shape overlap. The scores can be interpreted in terms agreement as slight [$0.11 \leq \text{score} < 0.2$], fair [$0.2 \leq \text{score} < 0.4$], moderate [$0.4 \leq \text{score} < 0.6$], substantial [$0.6 \leq \text{score} < 0.8$] and almost perfect [$0.8 \leq \text{score} \leq 1.0$]. Since such shape comparisons can be made only in the same space the set of comparisons are different from those in the left. The comparisons are classified into native-space comparisons (1 through 4) and the normalized-space comparisons (5 through 8). All the comparisons except 4th and 6th (dotted blue arrows) are made against the baseline either in native (BL) or normalized (NormBL) space. The 4th and 6th comparisons are made to measure if the dice and score distributions vary significantly from that of the 1st comparison, respectively.

For the comparisons involving the tracts obtained using combined data and those using individual repeat data (comparisons 1, 3, 5, 7) the dice and scores are estimated by bootstrapping the scores obtained by individual scans so that the scores reflect the inter-subject variance rather than inter-scan variance.

3. RESULTS

We first present the tract reconstructions for a representative subject in Fig. 4 which shows high visual consistency between the tract reconstructions showing that spatial normalization of tensors does not substantially effect the tractography based extractions of major white matter pathways such as the ones reconstructed here.

3.1 Consistency of Quantitative Summaries

Fig. 5 shows the BA plots for the comparisons 1 through 4 of Fig. 3(a), which compare the metrics (FA, MD and VOL) computed using normalized tensor volumes, while Fig. 6 shows the plots for the methods using native space metrics (comparisons 5 through 7 of Fig. 3(a).

The means (μ) and the standard deviations (σ) of the biases, the concordance coefficients (ρ) and Pearson correlation coefficients (r) (displayed in the respective titles of the scatter plots) are estimated using the measurements from all the tracts jointly. These plots reveal that the biases in the individual subjects are clustered around the means and fall within the 1.96 standard deviations showing that there is a high-level of agreement in the measurements. We can also clearly see that the bias shows no systematic dependence on the average for all the pathways except that the pathways with smaller average metrics tend to have larger variance in the bias. Hence when studying smaller pathways and smaller effect sizes, repeating the acquisitions and/or performing tracking in the native space are preferable.

To compare the seven different DTI processing streams in a simpler fashion we show the bar plots of their biases (averaged across subjects) in the measurements of FA, MD and log(VOL) of the 4 different tracts in Fig. 7. [1] We can first observe that overall biases in average FA, MD and log(VOL) of the tracts obtained by the seven methods compared to the baseline method are small and consistent across the structures, reflecting that the methods involving spatial transformations of tensors do not adversely effect the reconstruction of the major pathways such as the four considered here. The largest bias for FA is on the order of 0.07, it is 0.16 for MD and 0.25 for log(VOL). In case of VOL since the bias was estimated on the log-scale it reflects the ratio of the number of voxels rather than the difference. [2] We can then observe that the results are consistent across different tracts. This suggests that any of the seven methods involving spatial normalization of tensors does not confound statistical findings in population analyses if the effect sizes (population level variations) are above the variance in these bias levels. However if one is interested in analyzing effects on such an order, we can observe the following patterns which can be helpful in guiding the choice of appropriate DTI processing stream for population analysis.

FA Analyses—[1] We can observe that the normalized space methods (the first four bars in each plot) consistently underestimate more than the native space methods (the last three bars in each plot). This shows, as expected, that any sort of averaging (combining individual repeats) or interpolations (spatial transformations) is likely to underestimate (albeit only by small amount) resulting in a smoothing effect. Such an effect of registration has also been reported recently in (32). [2] The bias is smaller (~ 0.02) when performing tractography in the native space (comparisons 1, 4 and 6). The FA is consistently overestimated when using individual repeat data in the native space (comparison 6). This is consistent with empirically known fact that noise in the DWI data can inflate the FA of the tensors (33; 34). [3] Performing tractography and computing FA using the normalized tensor volumes underestimates the FA (comparisons 2 and 3) more than the other methods. [4] There is no significant difference between methods which compute one of the components (tractography or FA/MD/VOL) in the native space (comparisons 4 and 5).

Diffusivity and Volumetric Analyses—MD is relevant when analyzing diffusivities which reflect different biological properties than FA. VOL is used to measure the size of the tracts and is the total number of voxels in a tract. When studying the effects of the above two metrics, [1] we can observe that the estimate of the bias is small for all the tracts ($MD < 0.1$, $\log(VOL) < 0.25$), however for smaller pathways such as uncinate fasciculi, the variance in the bias is high (longer error bars) when using individual repeat data (comparisons 3 through 6) regardless of whether spatial normalization is involved in a processing stream. This implies that spatial normalization does not add undesired systematic variance in obtaining these metrics. [2] Further this indicates that when studying effect sizes < 0.2 for MD and VOL, it is better to acquire repeated acquisitions and estimate tensors using the combined data. [3] Interestingly we can also observe that in case of studying FA obtained via different

methods, the variances in the bias (errorbars) was smaller compared to those when concerned with MD and VOL. This is consistent with the fact that the diffusion tensor spatial normalization algorithms such as DTI-TK (12) minimize the distance metrics that emphasize the anisotropic components of the tensors.

3.2 Consistency of Geometric Shapes

Fig. 8 shows the shape consistencies for different tracts for the 8 different comparisons shown in Fig. 3(b). We can observe that all the methods show excellent shape overlaps reflecting that spatial normalization does not induce noise in the shapes of major tracts such as the ones reconstructed here. Although all tracts show consistent patterns in terms of shape agreement, CB and IFO have some of the lower scores. CB and IFO are more difficult pathways to localize in the brain. CB for example is adjacent to one of the largest tracts in the brain namely corpus callosum. The lower scores are hence attributable to the difficulty in their localization rather than the effects spatial transformations of tensors on tractography. From these scores we can observe that shape comparison scores are higher when tractography data are not warped. Also in the native space comparisons (comparison 5–8), the error bars are larger when comparing tracts in individual repeat data with those in combined data due to higher noise levels in the repeat data. Overall we can observe that combining the individual repeats and/or spatial normalization improves the shape overlaps due to smoothing effect.

Finally, Figs. 9 and 10 show the inter-subject and inter-scan variance of the dice coefficients and scores. The scores for CB and the IFO tracts are lower but as can be seen from the consistent score distributions in these figures, different methods involving spatial normalization of tensors do not induce any significant additional noise variance in the shapes of individual subject tracts.

4. CONCLUSIONS AND DISCUSSION

In this article we presented data showing the effects of diffeomorphic spatial transformation of DTI on tractography. The popular tensor reorientation strategies used in the spatial transformation of DTI, namely finite strain (FS) or preservation of principle direction (PPD) (11) although are adequately accurate for voxel based analyses (13) are not exact as in the case of rigid transformations. For example the results presented in (11) clearly show that even the averaged reorientation errors across an entire pathway are around 10° in highly coherent structures like corpus callosum. Hence it is likely that at the individual voxel level and for other complicated structures the reorientation errors are higher. Although some (global) tractography algorithmic frameworks (gibbs tracking) can be robust to such errors in the orientation information at the voxels most commonly used streamline methods are sensitive to local errors and can result in anatomically incorrect pathways. Although FS is more inaccurate than PPD, based on the premise that for diffeomorphic transformations with smaller Jacobian determinants FS is nearly as accurate with added computational advantages and is used in DTI-TK (12) which has been empirically shown to be best performing implementation for spatially normalizing DTI data (28; 35).

The data clearly indicate that non-linear and diffeomorphic transformations with PPD based tensor reorientation can indeed be adequately accurate for using tractography based reconstructions for population analyses. The tract reconstructions in the normalized tensors are mostly consistent with those in the native tensors. Large structures like corpus callosum the tract regions are more consistent when using normalized and native tensors compared to longer structures like cingulum and IFO. In such cases there are some spurious branches in the normalized tensors possibly due to smoothing effect and the worsened partial voluming of the ROIs used to filter the pathways. For smaller tracts that are relatively isolated such as

the uncinate fasciculus the consistency is higher. We would like to note that besides reorientation, another key component involved in DTI spatial transformation is tensor interpolation. It has been shown that the choice of metric for tensor interpolation (between Euclidean and log-Euclidean) can make a difference in the DTI metrics (36). However since log-Euclidean is shown to preserve the anisotropic components of the tensor which are important for tractography (37) we use log-Euclidean interpolation in our spatial transformations. Studying the effects of different interpolations on tractography is an interesting question but is not considered in this article. Finally our results show the effects of transformations that are a composite of rigid, affine and diffeomorphic. As discussed in (11) tensor reorientation can be perfect when using rigid-only transformations and diffeomorphic transformations are approximated using piecewise affine transformations. Hence we do not consider reporting individualized effects of rigid-only, affine-only transformations on tractography. Although an independent careful study might be required, the data here present assurance to the principle that appropriate local reorientations of diffusion models beyond DTI (such as fiber orientation distributions (38)) can be sufficient for non-local preservation of orientation information and faithful reconstruction of pathways.

Implications for Population Studies

When studying the WM substrates of various cognitive and psychopathological human behaviors, tractography data allows conducting statistically more powerful and computationally more efficient analyses. Although voxel-based analyses (VBA) of DTI scalar metrics are attractive, they are typically limited in power because of the need to address multiple comparisons problem in controlling type I error or false positive rate. Hence the frameworks such as tract-based spatial statistics (TBSS) became popular in WM studies. TBSS creates a skeleton on a population average FA map and condenses the local DTI properties onto the skeleton surface (39). TBSS is useful in reducing the number of comparisons; however, it does not necessarily correspond to specific tracts and the regional mapping of measures is based upon the local maximum FA, which may obscure focal effects.

Hence approaches such as region-of-interest (ROI) analyses which provide an attractive alternative for one to test hypotheses about *specific* white matter substrates have been developed and in particular those based on tractography such as tractography-guided spatial statistics (TGIS), tract based morphometry (TBM), tract specific analyses (TSA), tend to be provide more accurate localization of the pathways. In all of the above analytic frameworks it is of critical importance to obtain consistent measurements across individuals to minimize any confounding effects of image processing*. This is because diffusivity measures are highly sensitive to changes in the tissue microstructure. In particular, variance measures of the diffusivities (e.g., fractional anisotropy - FA) have become ubiquitous measures of differences in white matter associated with disease, injury, maturation and aging, or healthy variation. However, the diffusion anisotropy and other DTI measures are very heterogeneous across the brain - e.g., there is no uniform singular value or even range of FA in healthy white matter. Differences in the tissue microstructural features (i.e., axonal diameter and density, myelination, glial membranes) and white matter fiber directional organization (i.e., fiber crossing, dispersion, bending) will influence the diffusion measurements regionally, though it is generally assumed that the structural organization is consistent across subjects and therefore regions may be compared across individuals. Manually defined regions of interest can achieve high spatial specificity, but are sensitive to inconsistencies in the regional definition. Improvements to spatial normalization have also enabled the application

---

*This is a different type of control than controlling type I error.

of regional templates for automated region-of-interest analyses either in the native or normalized spaces (40). Although spatial normalization is a widely accepted computational framework for establishing correspondences across individuals, the analyses are sensitive to the accuracy of the spatial normalization and the amount of smoothing introduced by the image processing to generate the statistical maps. These smoothing effects may be compensated to a certain degree (41) though they do not eliminate the effects of undesirable partial volume averaging. To our best knowledge, it has not been verified rigorously, prior to this work, if spatial normalization, besides inducing such smoothing effects, effects tractography results which are sensitive to accumulated errors. Nevertheless tractography has been used to define tract-specific regions in DTI data. This is because although this might introduce differences in the tract reconstructions there is no reason to expect systematic subject specific variance and hence is viable for population studies.

Tractography can facilitate more accurate localization of ROIs in the WM and can be identified by following protocols such as (2; 3). Following such protocols in each individual subject data would be cumbersome in a population study. Hence often the “filter” ROIs namely, waypoint and exclusion ROIs, needed to identify the WM pathways are traced on a population specific atlas and then inverse warped into the subject space (42). It has been indicated in (40) tract reconstructions using DTI can only be “validated” using strong *a priori* knowledge about the shapes of the tracts. Given such a limitation, the identification of filter ROIs in the atlas can be more efficient when tractography data are available in the atlas which can provide instant feedback on the verification of the *a priori* shapes of the tracts. Population specific atlases benefit from averaging across many subjects in the normalized space and hence any the minor reorientation errors at the individual voxel level can be expected to cancel out. Hence tractography on the atlas is performed in applications which perform tract specific statistical analyses (8; 15; 16). These applications inverse warp not the filter ROIs but the binarized tractography masks (8; 15) or fiber tracts themselves (16) identified on the atlas.

Although there have been applications, such as above, that perform tractography on the atlas there are only a few articles using tractography in the normalized space for statistical analyses. These applications do not use the shape of the tracts but rather their count between pairs of regions for modeling structural brain networks in the normalized space (9). Many approaches such as tractography-guided spatial statistics (TGIS) (17) and tract based morphometry (TBM) (18) however still prefer tracking in the native space. Our data which shows that spatial transformation of DTI preserves the shapes of the tracts implies that (17) can obtain the 3-D parametric meshes of the tracts by using normalized space tractography and thus achieve computational efficiency. In (18) for example, one can avoid warping of the tracts and an ill-posed tract clustering step since performing normalized space tractography can establish correspondences across subjects needed for population analyses. One can also imagine building into prior information into tractography algorithms when tracking in normalized space. It might be an interesting application to generating voxel-wise connectivity maps using tractography in normalized space. Without performing tracking in normalized space such an application would be extremely hard to implement.

Our results quantitatively demonstrate that normalized space tractography produces anatomically consistent structures compared to native space tractography, thus providing a firm ground for using tractography data in population analyses. Both macrostructural and microstructural properties of tractography results are preserved during spatial normalization. This has two important implications for population analyses: [1] except when studying extremely small effects, one can safely extract quantitative properties from the tractography data in the normalized space for performing ROI as well as tract specific shape analyses, [2] the high-dimensional non-linear spatial transformations of the diffusion tensor imaging data

preserve non-local anatomical connections of the white matter. Hence one can also build network models using normalized space tractography (e.g. (9, 10)) for studying topological differences between individuals and groups.

Acknowledgments

We thank Frances Haeberli and Chad M. Ennis for assisting in the data collection and Steve R. Keckemeter for helpful discussions. This work was supported by the NIMH R01 MH080826 and R01 MH084795 and the NIH Mental Retardation/Developmental Disabilities Research Center (MRDDRC Waisman Center), NIMH 62015, the Autism Society of Southwestern Wisconsin, the NCCAM P01 AT004952-04 and the Waisman Core grant P30 HD003352-45.

References

- [1]. Alexander A, Hurley S, Samsonov A, Adluru N, et al. Characterization of cerebral white matter properties using quantitative magnetic resonance imaging stains. *Brain Connectivity*. 2012; 1(6): 423–446. [PubMed: 22432902]
- [2]. Wakana S, Jiang H, Nagae-Poetscher L, van Zijl P, Mori S. Fiber tract based atlas of human white matter anatomy. *Radiology*. 2004; 230:77–87. [PubMed: 14645885]
- [3]. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*. 2008; 44(8):1105–1132. [PubMed: 18619589]
- [4]. Arfanakis K, Gui M, Lazar M. Optimization of white matter tractography for pre-surgical planning and image-guided surgery. *Oncol Rep*. 2006; 15:1061–1064. [PubMed: 16525701]
- [5]. Calamante F, Tournier J, Smith R, A, C. A generalised framework for super-resolution track-weighted imaging. *NeuroImage*. 2012; 59(3):2494–2503. [PubMed: 21925280]
- [6]. Jbabdi S, Johansen-Berg H. Tractography: where do we go from here? *Brain Conn*. 2011; 1(3): 169–183.
- [7]. Adluru N, Hinrichs C, Chung M, JE. L, et al. Classification in DTI using shapes of white matter tracts. *IEEE Engineering in Medicine and Biology Society*. 2009:2719–2722.
- [8]. Zhang H, Awate S, Das S, Woo J, et al. A tract-specific framework for white matter morphometry combining macroscopic and microscopic tract features. *Medical Image Analysis*. 2010; 14(5): 666–673. [PubMed: 20547469]
- [9]. Chung M, Adluru N, Dalton K, Alexander A, Davidson R. Scalable brain network construction on white matter fibers. *SPIE Medical Imaging*. 2011
- [10]. Adluru N, Chung MK, Lange NT, Lainhart JE, Alexander AL. Applications of epsilon radial networks in neuroimage analyses. *Adv. in Image and Video Technology - LNCS*. 2012:236–247.
- [11]. Alexander D, Pierpaoli C, Basser P, Gee J. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Trans. Med. Imaging*. 2001; 20:1131–1139. [PubMed: 11700739]
- [12]. Zhang H, Yushkevich P, Alexander D, Gee J. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Medical Image Analysis*. 2006; 10:764–785. [PubMed: 16899392]
- [13]. Zhang H, Avants B, Yushkevich P, Woo J, et al. High-dim. spatial norm. of DTI improves the detection of WM differences: An ex. study using ALS. *IEEE Trans. Med. Imaging*. 2007; 26:1585–1597. [PubMed: 18041273]
- [14]. Lazar M, Alexander A. An error analysis of white matter tractography methods: synthetic diffusion tensor field simulations. *NeuroImage*. 2003; 20(2):1140–1153. [PubMed: 14568483]
- [15]. Yushkevich P, Zhang H, Simon T, Gee J. Structure-specific statistical mapping of white matter tracts. *NeuroImage*. 2008; 41(2):448–461. [PubMed: 18407524]
- [16]. Goodlett C, Fletcher T, Gilmore J, Gerig G. Group analysis of DTI fiber tract statistics with application to neurodevelopment. *NeuroImage*. 2009; 45(1):S133–S142. [PubMed: 19059345]
- [17]. Jungsu SO, InChan S, JaeSung L, Hyejin K, et al. TGIS in DTI for the detection of gender diff. of fiber integrity in the midsag. and parasag. corpora callosa. *NeuroImage*. 2007; 36(3):606–616. [PubMed: 17481923]

- [18]. O'Donnel L, Westin C, Golby A. Tract-based morph. for WM group anal. *NeuroImage*. 2009; 45:832–844. [PubMed: 19154790]
- [19]. Besseling R, Jansen J, Overvliet G, Vaessen M, et al. Tract specific reproducibility of tractography based morphology and diffusion metrics. *PLoS ONE*. 2012; 7(4):e34125. [PubMed: 22485157]
- [20]. Ziyen U, Sabuncu M, Lauren J, Westin C. Nonlinear registration of diffusion MR images based on fiber bundles. *MICCAI LNCS*. 2007; 4791(1):351–358.
- [21]. Zöllei L, Stevens A, Huber K, Kakunoori S, B, F. Improved tractography alignment using combined volumetric and surface registration. *Neuroimage*. 2010; 51(1):206–213. [PubMed: 20153833]
- [22]. Zvitia O, Mayer A, Shadmi R, Miron S, Greenspan H. Co-registration of WM tractographies by adaptive-mean-shift and GMM. *IEEE Trans. Med. Imaging*. 2010:132–145. [PubMed: 19709970]
- [23]. Heiervang E, Behrens T, Mackay C, Robson M, et al. Between session reproducibility and between subject variability of diffusion MR and tractography measures. *NeuroImage*. 2006; 33(3):867–877. [PubMed: 17000119]
- [24]. Adluru, N. and et al. Eval. tract. in the spat. norm. DTI data. *ISMRM*. 2012
- [25]. Adluru N, et al. Spat. norm. of DTI preserves tract recon. of major WM pathways. *HBM*. 2012; 401
- [26]. Smith S, et al. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*. 2004; 23:208–219.
- [27]. Cook P, et al. Camino: Open-source DMRI recon. and proc. *ISMRM*. 2006; 2759
- [28]. Adluru N, Zhang H, Fox AS, et al. A diffusion tensor brain template for rhesus macaques. *NeuroImage*. 2012; 59(1):306–318. [PubMed: 21803162]
- [29]. Wang, R. and et al. Diffusion toolkit: A software package for diffusion imaging data proc. and tractography. *ISMRM*. 2007; 15:3720.
- [30]. Altman D, Bland J. Measurement in medicine: the analysis of method comparison studies. *Statistician*. 1983; 32:307–317.
- [31]. Landis J, Koch G. The measurement of obs. agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- [32]. Bockhorst K, et al. Effect of reg. on FA values. *ISMRM*. 2012; 20:3556.
- [33]. Pierpaoli C, Basser P. Toward a quantitative assessment of diffusion anisotropy. *Magn. Reson. Med*. 1996; 36:893–906. [PubMed: 8946355]
- [34]. Jones D, Basser P. Squashing peanuts and smashing pumpkins: how noise distorts diffusion-weighted MR data. *Magn. Reson. Med*. 2004; 52:979–993. [PubMed: 15508154]
- [35]. Wang Y, et al. DTI registration in atlas based fiber analysis of infantile krabbe disease. *NeuroImage*. 2011; 55:1577–1586. [PubMed: 21256236]
- [36]. Pasternak O, Sochen N, Basser P. The effect of metric selection on the analysis of diffusion tensor MRI data. *NeuroImage*. 2010; 49:2190–2204. [PubMed: 19879947]
- [37]. Fillard P, Arsigny V, Pennec X, Ayache N. Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics. *IEEE Trans. Med. Imag*. 2007; 26:1472–1482.
- [38]. Raffelt D, Tournier D, et al. Symm. diffeo. registration of FODs. *NeuroImage*. 2011; 56(3):1171–1180. [PubMed: 21316463]
- [39]. Smith S, Jenkinson M, Johansen-Berg H, Rueckert D, et al. Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006; 31:1487–1505. [PubMed: 16624579]
- [40]. Mori S, Oishi K, Faria A. White matter atlases based on diffusion tensor imaging. *Current Opinions in Neurology*. 2009; 22(4):362–369.
- [41]. Lee J, Chung M, Lazar M, DuBray M, et al. A study of diffusion tensor imaging by tissue-specific, smoothing-compensated voxel-based analysis. *NeuroImage*. 2009; 44(3):870–883. [PubMed: 18976713]

- [42]. Pagani E, et al. A method for obtaining tract-specific DT-MRI measurements in the presence of disease: appl. to patients with clinically isol. synd. suggestive of MS. *NeuroImage*. 2005; 26(1): 258–265. [PubMed: 15862226]

Figure 1.

(a) The two different types of tensor estimation performed in the study. s_i^j : For each subject, 8 b_0 s and 48 DWI volumes are used in tensor volume estimation for each of the four individual repeats. The motion correction step is performed for each repeat independently by registering all the images to the corresponding volume's first b_0 image. $\bar{s}_i$: The four repeats (32 b_0 s and 192 DWI volumes) are combined by averaging resulting in a higher-quality tensor estimation. Here the motion correction is performed by registering all the volumes of all the repeats to the first b_0 of the first repeat. (b) The spatial transformations estimated using DTI-TK. The 16 different tensor volumes (4 individual repeats for 4 different subjects) and the 4 combined tensor volumes are registered to a template obtained from a different study with the same acquisition protocol. The set of transformations (F) are stored and used in the creation of the study specific atlas ($\mathcal{A}$) as shown in Eq. (2). (c) The key schematic illustrating the eight different processing streams for each subject. The accompanying Table 1 describes the various symbols used in this figure.

Figure 2.

Sample viewpoints of the reconstruction of the four tracts using waypoint and exclusion filter regions (FR). (a) Corpus callosum (CC). (b) Cingulum bundle (CB). (c) Inferior fronto-occipital fasciculus (IFO). (d) Uncinate fasciculus (UF).

Figure 3.

Graphical summary of different sets of comparisons. (a) Seven different methods are compared against the baseline method to measure the agreement of average FA, MD and VOL of the specific tracts. (b) The pairs of comparisons for measuring the macrostructural consistency using dice coefficients and scores.

Figure 4.

Visual consistency of the tract reconstructions using native ($\bar{s}_i$) and spatially normalized tensors ($F_i(\bar{s}_i)$). The red circles highlight some minor inconsistencies between the reconstructions of cingulum bundle (CB) and the inferior fronto-occipital (IFO).

Figure 5.

Bland-Altman plots to measure the agreements between the average fractional anisotropy (FA) (top), mean diffusivity (MD) (middle) and log-volume (log(VOL)) (bottom) of the four different white matter tracts. The columns corresponds to the comparisons 1–4 in Fig. 3(a) which evaluate agreements between normalized-space measurements and baseline (BL) method. Although the overall bias in all the four methods is small, they consistently (and only slightly) underestimate the measurements.

Figure 6.
Similar to Fig. 5 but for the comparisons 5–7 of Fig. 3(a).

Figure 7.

Bar-plots with error-bars showing the bias of the seven different methods (corresponding to those in Fig. 3(a)) of extracting the quantitative summaries of the four tracts compared to the baseline method ($\bar{\mu}(\mathcal{T}_*(\bar{s}))$).

Figure 8.

Comparison of different processing methods presented in Fig. 3(b). Each group of bars represents dice coefficient (left) and scores (right). Different structures are shown in different colors. The first four groups of bars reflect the comparisons in the normalized space and the last four in the native space.

Figure 9. Inter-subject and inter-scan variance in the v scores for the eight different comparisons in Fig. 3(b). The top row shows the comparisons in the normalized space (comparisons 1 through 4) and the bottom four in the native space (comparisons 5 through 8).

Figure 10.
Similar to Fig. 8 but shows the results using dice similarity coefficients.

Table 1

Summary of different symbols presented in Fig. 1. We would like to note that F acts on tensors in the native space while on pointsets in the normalized space. For example, to warp $\bar{s}_i$ to the normalized space, we apply F_i . On the other hand, applying F_i to tracts (which are pointsets) transforms them from normalized space to native space. Furthermore, although $\{F_i^j\}_{j=1}^4$ and F_i are expected to be identical they can be slightly different based on the amount of motion between the repeat scans.

Symbol	Description
$\bar{s}_i, s_i^j$	Tensor volume for subject i using combined DWI and j^{th} repeat respectively.
F_i, F_i^j	Smooth invertible diffeomorphic transformations for $\bar{s}_i$ and s_i^j respectively.
$A \equiv \mathcal{T}_*(\bar{s}_i)$	Tractography based reconstructions using combined data in the native space for subject i .
$A' \equiv \mathcal{T}_*(F_i(\bar{s}_i))$	Tractography based reconstructions using the spatially normalized combined data for subject i .
$B \equiv \mathcal{T}_*(s_i^j)$	Tractography based reconstructions using j^{th} scan of the i^{th} subject.
$B' \equiv \mathcal{T}_*(F_i^j(s_i^j))$	Tractography based reconstructions using spatially normalized j^{th} scan of the i^{th} subject.
$F_i^{-1}(A)$	Warped A into the normalized space and inverse warped A into the native space respectively.
$F_i(A')$	Inverse warped A into the native space respectively.
$F_i^{j-1}(B)$	Warped B into the normalized space and inverse warped B into the native space respectively.
$F_i^j(B')$	Inverse warped B into the native space respectively.
$\bar{\mu}(\mathcal{T})$	Quantitative summary of FA/MD/VOL by binarizing the rasterized tractography $\mathcal{T}$.