Published in final edited form as:
Neuroimage. 2010 February 15; 49(4): 3047–3056. doi:10.1016/j.neuroimage.2009.11.043.

Automated vs. Conventional Tractography in Multiple Sclerosis: Variability and Correlation with Disability

Daniel S. Reich, Ph.D., M.D.[1,2,3,4], **Arzu Ozturk, M.D.**[1], **Peter A. Calabresi, M.D.**[2], and **Susumu Mori, Ph.D.**[1,3]

[1]Department of Radiology, Johns Hopkins University, 600 N Wolfe St, Baltimore, MD 21287.

[2]Department of Neurology, Johns Hopkins University, 600 N Wolfe St, Baltimore, MD 21287.

[3]F. M. Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute, 707 N Broadway, Baltimore, MD 21205.

Abstract

Diffusion-tensor-imaging fiber tractography enables interrogation of brain white matter tracts that subserve different functions. However, tract reconstruction can be labor and time intensive and can yield variable results that may reduce the power to link imaging abnormalities with disability. Automated segmentation of these tracts would help make tract-specific imaging clinically useful, but implementation of such segmentation is problematic in the presence of diseases that alter brain structure. In this work, we investigated an automated tract-probability-mapping scheme and applied it to multiple sclerosis, comparing the results to those derived from conventional tractography. We found that the automated method has consistently lower scan-rescan variability (typically 0.7%−1.5% vs. up to 3% for conventional tractography) and avoids problems related to tractography failures within and around lesions. In the corpus callosum, optic radiation, and corticospinal tract, tract-specific MRI indices calculated by the two methods were moderately to strongly correlated, though systematic, tract-specific differences were present. In these tracts, the two methods also yielded similar correlation coefficients relating tract-specific MRI indices to clinical disability scores. In the optic tract, the automated method failed. With judicious application, therefore, the automated method may be useful for studies that investigate the relationship between imaging findings and clinical outcomes in disease.

Keywords

diffusion tensor imaging; tractography; magnetization transfer imaging; multiple sclerosis; corticospinal tract; corpus callosum; visual system

4Current affiliation and address for correspondence: Translational Neuroradiology Unit, Neuroimmunology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health; 10 Center Drive, MSC 1400; Building 10, Room 5C103; Bethesda, MD 20892; reichds@ninds.nih.gov.

Prior presentation: Portions of this work were presented at the 2009 annual meeting of the International Society for Magnetic Resonance in Medicine (Honolulu, Hawaii: April 18–24, 2009).

Introduction

In diseases such as multiple sclerosis (MS), amyotrophic lateral sclerosis, and stroke, interrogation of specific white matter tracts, commonly performed with tractography algorithms based on diffusion tensor imaging (DTI) (Basser et al., 2000; Conturo et al., 1999; Mori et al., 1999; Parker et al., 2003), has allowed noninvasive monitoring and improved understanding of the links between white matter structure and disability – demonstrating, for example, that motor disability is linked with corticospinal-tract imaging abnormalities (Lee et al., 2005; Reich et al., 2008) and impaired cognition with damage to the corpus callosum (Lin et al., 2008).

However, with many of the commonly used tractography algorithms, tract reconstruction is laborious, time consuming, and often highly variable from scan to scan, factors that markedly diminish its routine use in the clinic. Variability in reconstructed tract volume is particularly problematic, but even variability in tract-specific MRI indices, which is considerably smaller, can limit sensitivity for detecting changes over time that may cause or correlate with disability (Heiervang et al., 2006; Reich et al., 2006; Wakana et al., 2007). These difficulties may be compounded in disease, where disruption in tissue microstructure may interfere with the features of the diffusion tensor that allow reliable delineation of tracts (Pagani et al., 2007a; Wheeler-Kingshott and Cercignani, 2009). For all of these reasons, the development of rapid, automated, and reproducible methods for tract-specific imaging will go a long way toward improving the utility of tractography in the clinic and in high-throughput clinical trials.

At least two promising solutions to this problem have been proposed. One of these, so-called tract-based spatial statistics (Smith et al., 2006), allows comparison of the imaging features of different tracts across individuals by reducing those tracts to their cores or “skeletons,” a maneuver that greatly simplifies the registration problem. This method is useful for performing whole-brain analysis, but in its original implementation it does not separate the brain into distinct, functionally relevant tracts. An additional drawback of this method is that it is based on measurement of tissue anisotropy, which may be severely disrupted in disease.

A second solution that is receiving increasing attention uses atlases of varying complexity to segment the brain into functional components. Along these lines, we recently introduced a method for creating tract-probability maps from DTI data (Hua et al., 2008). This method is based on identifying the probability with which voxels in spatially coregistered brain images are identified as belonging to a particular tract. In essence, the tract of interest is reconstructed from the DTI data in many individuals, and the number of subjects in which each voxel is included in the reconstructed tract is tabulated. This leads to a tract-probability map, which can then be multiplied by an arbitrary set of coregistered brain images. MRI indices along the tract of interest, derived either from the DTI data or from coregistered, quantitative scans of other types, can thus be weighted by the probability that each voxel belongs to the tract.

This method offers the promise of rapid, automatic evaluation of tract-specific MRI indices without the need for dedicated tract reconstruction in every individual and without concern for tractography failures in diseased regions. If successful, tract-specific imaging could thus be made a routine part of clinical practice, giving physicians the option to readily image functionally relevant portions of the brain in individuals with disease.

The success of any atlas-based method is determined by the quality of the registration between the test and reference scans. Perfect registration is difficult to achieve even in normal brains due to variability in the position and size of the gyri, sulci, and ventricles; the problem is compounded in diseased brains (Pagani et al., 2007a; Pagani et al., 2007b). One obvious example is a space-occupying lesion, such as a tumor, that causes mass effect and architectural distortion in the brain. This problem exists even in the absence of mass lesions, because there

is no obvious registration target for areas of abnormal signal intensity (such as MS lesions) and because many diseases (including MS) are characterized brain atrophy, the extent of which may vary across the brain. In our initial paper (Hua et al., 2008), we presented a single case example of a hemiparetic individual with MS and determined that a tract-probability-map approach identified abnormalities in the contralateral intracranial corticospinal tract. However, a more thorough investigation is clearly needed to determine whether such an approach can be applied more widely. This paper represents a step in that direction.

Since conventional tractography is currently at best an imperfect technique, subject to intrapersonal, interpersonal, and implementation-dependent variability (Burgel et al., 2009), there is no gold standard by which to judge successful application of tract-probability maps to patients. Nonetheless, it is important to compare the results of both approaches in the case of disease, with several goals: (1) to assess the extent to which each tract is successfully identified by the tract probability approach – for example, whether the identified tract position is non-anatomic; (2) to determine whether the tract-probability-map method, compared to conventional tractography, results in systematic changes in various MRI indices; (3) to assess whether the discrepancies between conventional tractography and tract-probability mapping differ across tracts; (4) to determine whether the association between disability and tract-specific imaging results is hampered, preserved, or improved when switching from one approach to the other; and (5) to measure the scan-rescan variability in results derived from the two methods.

In this paper, we address these issues in the context of MS, a demyelinating and neurodegenerative disease of the central nervous system in which tract-specific imaging may be especially important. This is because of the poor correlation between disability scores and whole-brain MRI abnormalities, such as T2 lesion number and volume (Goodin, 2006). Tract-specific imaging can focus attention on arbitrary, functionally relevant tracts (Mori et al., 2008) and has the potential to improve these correlations and thus provide an objective measure of the status of a patient's disease. At the same time, MS presents a challenge for the tract-probability-map approach because of the presence of lesions and prominent brain atrophy, processes that may limit the degree to which diseased brains can be registered to the normal atlas.

Methods

The method for creating tract probability maps that we use here is based on, but modified from, our prior work (Hua et al., 2008). The major conceptual difference is that rather than transforming all brains to a standard space, such as the Montreal Neurological Institute atlas, we used a DTI scan from a single healthy individual as our reference. Our method is shown graphically in Figure 1, and a detailed description follows.

MRI protocol

MRI scans at 3T (Philips) were obtained with Institutional-Review-Board approval and following signed, informed consent of the study participants. We used the body coil for transmission and either a 6-channel head coil or the 8 head elements of a 16-channel neurovascular coil for reception. For each individual, we obtained a pair of DTI scans with 2.2 mm isotropic voxels and the following scan parameters: TE, 69 ms; TR, automatically calculated ("shortest"); slices, 60 or 70; SENSE factor, 2.5; non-collinear diffusion directions, 32 (Philips "overplus" scheme); high *b*-value, 700 s/mm^2; low *b*-value ("b_0"), approximately 33 s/mm^2; repetitions, 2. We also performed a double-echo fast-spin-echo sequence to obtain proton-density- and T2-weighted scans (acquired resolution: 1.1 × 1.1 × 2.2 mm; TE: 28.2 ms and 90 ms; TR: 4162 ms; SENSE factor: 2; repetitions: 1) as well as a magnetization-transfer

sequence (acquired resolution: 1.5 × 1.5 × 2.2 mm; TE: 15 ms; TR: 64 ms; SENSE factor: 2; MT pulse: sinc-shaped, 1.5 kHz off-resonance; repetitions: 3).

After removing the data to a separate workstation, we estimated T2-relaxation-time (qT2) and magnetization-transfer-ratio (MTR) maps as previously described (Reich et al., 2006). We then coregistered all scans (diffusion-weighted, qT2, and MTR) to the minimally diffusion-weighted scan from the first DTI sequence using a 6-parameter, rigid-body transformation implemented in the AIR program (Woods et al., 1998). We next used CATNAP (http://iacl.ece.jhu.edu/~bennett/catnap/catnap.shtml) (Landman et al., 2007) to estimate the diffusion tensor and create maps of fractional anisotropy (FA), mean diffusivity (MD), parallel diffusivity ($\lambda_{\parallel}$), and perpendicular diffusivity ($\lambda_{\perp}$). These four quantities, together with qT2 and MTR, are hereafter termed *MRI indices*.

Tractography

In every scan, we used the fiber assignment by continuous tractography method (Mori et al., 1999) to reconstruct the courses of the corpus callosum, corticospinal tracts, optic radiations, and optic tracts and saved the results as binary masks. To reconstruct the corpus callosum, we placed three regions of interest (ROI) around the tract on sagittal reformations of the axially acquired data (one in the midline and one each at the lateral borders of the lateral ventricles as visualized on coronal reformations at the midpoint of the corpus callosum) and required reconstructed tracts to pass through these ROIs (Ozturk et al., 2009). For the corticospinal tract, we placed three ROIs around the tract on axial slices: one in the rostral medulla, one in the rostral pons, and one at the level of the precentral gyrus (Reich et al., 2006). For the optic radiation, we placed two ROIs around the tract on coronal reformations: one at the level of the lateral geniculate nucleus and the second at the level of the subcortical white matter in the occipital lobe (Reich et al., 2009). Finally, for the optic tract, we placed two ROIs around the tract on coronal reformations: one at the level of the optic chiasm and the second at, or just anterior to, the lateral geniculate nucleus (Dasenbrock et al., 2009).

We used an FA threshold of 0.13 as the starting and stopping criteria for reconstructed fibers and a voxel-to-voxel turning-angle threshold of 30 degrees (except for the corticospinal tract, where the turning-angle threshold was 40 degrees). We analyzed only the portions of the fibers that ran between the two outermost ROIs and excluded from consideration the portions that passed beyond those ROIs.

Tract probability maps

Via a 12-parameter affine transformation model implemented in FLIRT (http://www.fmrib.ox.ac.uk/fsl/flirt) (Jenkinson et al., 2002), we coregistered the b_0 scans for each of remaining 26 healthy volunteers to a single “reference” scan (35-year-old healthy woman) that was chosen for its excellent image quality. We used the transformation matrix derived from this registration to register the binary tract masks to the reference space. In this reference space, we estimated the probability that each tract traverses each voxel by calculating the voxel-wise average of the registered tract masks. We removed from consideration all voxels with probability <0.05, as these voxels were most likely to be spurious. These probability values were used to weight the average MRI indices for each tract, as described in the next paragraph. Figure 2 shows the four tracts considered here superimposed on two selected axial sections from the minimally diffusion-weighted map of the reference scan. Note that these tract-probability maps tend to be larger than the true tracts, particularly for small tracts such as the optic tract.

Tract-specific MRI indices

To measure tract-specific MRI indices for a test scan, we transformed the minimally diffusion-weighted scan from that scan into the reference space, again using a 12-parameter transformation implemented in FLIRT. We used the resulting transformation to coregister the quantitative MRI maps (FA, MD, $\lambda_{\parallel}$, $\lambda_{\perp}$, qT2, and MTR) that had already been registered to the b_0 acquisition for that scan (see above). To restrict our attention to white matter tracts and to reduce the likelihood of contamination by cerebrospinal-fluid and gray matter voxels due to imperfect registration between the test and reference scans, we removed all voxels with FA <0.25. We then used each of the seven tract masks (corpus callosum and bilateral corticospinal tract, optic radiation, and optic tract) to calculate weighted, tract-specific averages for each quantitative MRI map.

To quantify the degree of overlap between the reconstructed tracts in the test and reference scans, we summed the weights for each of the surviving voxels in the test scan and divided the result by the summed weights for all voxels in the atlas. The resulting *overlap index* can range from 0 (no overlap) to 1 (complete overlap) and tends to be higher if more high-probability voxels (i.e., those with weights close to 1) are retained in the test scan after registration and masking.

Spatially normalized tract profiles

To visualize the spatial variation in MRI indices along each tract, we created spatially normalized tract profiles (Reich et al., 2007). In the reference space, we divided each tract into segments and subsegments, where the number of subsegments was chosen so that the size of each was approximately 1 mm. For the optic tract, optic radiation, and corticospinal tract, these segments were placed along the rough anatomical course of the tract. Thus, we divided the optic tract into 40 subsegments along the anteroposterior axis, the optic radiation into 60 subsegments along the same axis, and the corticospinal tract into 6 segments, each containing 20 subsegments, along the rostrocaudal axis. The boundaries of the 6 segments were placed at predetermined anatomical locations as described previously (Reich et al., 2007). Because of the distinctive morphology of the corpus callosum, we divided it into 6 segments along the anteroposterior axis (rather than along the mediolateral orientation of its component fibers) according to a previously published scheme (Witelson, 1989). Each of these segments – rostrum and genu together, rostral body, anterior midbody, posterior midbody, isthmus, and splenium – was divided into 20 equal subsegments.

To create the tract profile for a given MRI index, we analyzed the MRI-index maps after registration to the atlas space. We calculated the average value of the index within sliding windows, 5 subsegments wide, centered on each subsegment. We then plotted these average values as a function of tract position (anteroposterior for the optic tract, optic radiation, and corpus callosum, and rostrocaudal for the corticospinal tract). Within each subsegment, we calculated the mean value and 95% confidence limits across the 27 healthy volunteers, assuming a normal distribution of values.

Conventional tractography

We compared the results of the above procedures, which generate tract-specific MRI indices and spatially normalized tract profiles via a tract-probability-map method, to the results obtained from conventional tractography. Specifically, we performed tractography separately for each scan without transformation to the reference space. From these data, we recorded median tract-specific indices and constructed spatially normalized tract profiles. The method used to construct the tract profiles was identical to the one described above, except that the segment boundaries were defined in the native space rather than in the reference space.

Brain volume

We used the DTI images to estimate the brain parenchymal fraction. To do this, we first removed the skull using BET (http://www.fmrib.ox.ac.uk/fsl/bet2) (Smith, 2002). We then used an affine transformation in AIR (Woods et al., 1998) to register a volume consisting of the average of all diffusion-weighted images to the corresponding volume of a single-subject DTI atlas ("JHU_MNI_SS") that is available for download (http://www.mristudio.org). This atlas, which contains 1 mm isotropic voxels in a field-of-view of 181 × 217 × 181 mm, has been registered to the standard MNI-152 atlas. We then selected for analysis a portion of the brain between the mammillary bodies inferiorly and the centrum semiovale superiorly (coordinates $54 \leq z \leq 124$) and removed a small section of the anterior frontal lobes ($y \leq 37$) that is particularly susceptible to echo-planar-imaging-related distortion. We segmented the cerebrospinal fluid by selecting voxels with MD ≥ 1.7 μm^2/ms. We used the resulting brain parenchymal and cerebrospinal-fluid masks to calculate the brain parenchymal fraction, defined as the ratio of the volume of brain parenchyma to the total volume of brain and cerebrospinal fluid within the area of interest.

Disability scores

Our previous work (and that of others) has demonstrated a link between MRI indices in the corpus callosum and cognition, assessed using the 3-second version of the Paced Auditory Serial-Addition Test (PASAT) (Ozturk et al., 2009) and between MRI indices in the optic radiation and visual acuity at 1.25% contrast (Reich et al., 2009). As a clinical surrogate for the other tracts, we also used 1.25%-contrast visual acuity for the optic tract and the 25-foot walking time – which, like the PASAT, is a component of the Multiple Sclerosis Functional Composite (Cutter et al., 1999) – for the corticospinal tract. We included in the regression analysis the MS scans for which there was a corresponding disability measurement within 30 days.

Statistics

All statistical analyses were performed in Stata 9.0 (Stata Corporation, College Station, TX). Because we had multiple scans per participant, we used mixed-effects linear regression models with random (participant-specific) intercepts to assess the differences in overlap indices between healthy volunteers and individuals with MS as well as the degree to which tract-specific MRI indices calculated by the tract-probability-mapping method can predict tract-specific MRI indices calculated via conventional tractography. Such models appropriately account for different numbers of scans per participant, and the formulation assumes that the relationship between MRI indices derived from the two methods does not change over time.

We ran the models separately for each MRI index. Each model was of the form:

$$C_{ij}=\beta_0+\beta_1 T_{ij}+\beta_2 d_i+\beta_3 d_i T_{ij}+u_i+\varepsilon_{ij} \quad (1)$$

where C_{ij} and T_{ij} are, respectively, the tract-specific MRI indices for participant i at scan j, calculated using the conventional and tract-probability-map methods; d_i is a categorical variable defining the presence (d_i=1) or absence (d_i=0) of MS; u_i is a participant-specific random intercept; and ε_{ij} is the residual. We fit the coefficients β_k using generalized-least-squares estimators. The strength of the association between conventional and tract-probability-map results was gauged using an overall R^2 value that takes into account both within-participant and between-participant effects. The slopes of the relationship between MRI indices from the two methods are β_1 for healthy volunteers and $\beta_1+\beta_3$ for participants with MS.

To measure the reproducibility over time of tract-specific MRI indices, we calculated the average within-participant coefficients of variation (standard deviation divided by mean) for both methods. The goal of this paper is not to describe the links between disability scores and tract-specific MRI indices *per se*, but rather to determine whether the conventional and tract-probability-map approaches yield similar results. Thus, to assess the relationship between disability scores and tract-specific MRI indices in the MS group, we calculated Pearson correlation coefficients without adjustments for multiple scans per participant or covariates such as participant age, sex, disease duration, or brain atrophy. Based on prior work, we expected the associations between our four disability scores and the corresponding tract-specific MRI indices to be modest at best.

Results

We obtained 45 scans from 27 healthy volunteers (age, mean ± standard deviation: 33.2 ± 17.2 years) and 244 scans from 88 individuals with MS (46.4 ± 11.7 years). In the MS group, 50 individuals had the relapsing-remitting form of the disease, 24 had the secondary-progressive form, and 14 individuals had the primary-progressive form. The median expanded disability status scale (EDSS) (Kurtzke, 1983) score was 3.5 (range: 0–8), and the mean disease duration, calculated from the onset of subjective symptoms, was 10.7 years (standard deviation: 8.9). The brain parenchymal fraction was 0.908 ± 0.038 in the healthy-volunteer group and 0.846 ± 0.049 in the MS group ($p<0.001$ adjusting for age).

Figure 3 shows the extent of overlap (mean and standard deviation) between tract masks derived from tract-probability maps and the corresponding atlas tract. As described in the Methods section, the overlap index can range from 0 to 1, with 1 indicating perfect overlap. Overlap indices were significantly higher in the healthy-volunteer group for all tracts even after accounting for differences in age between the two groups, indicating better registration to the atlas in the healthy-volunteer group – not a surprising finding given that the atlas was itself derived from the healthy volunteers. Overlap was best in the corticospinal tract and optic radiation and worst in the optic tract and corpus callosum.

Figure 4 compares MRI indices obtained from the two tract-reconstruction methods. MRI indices obtained from conventional tractography are plotted along the vertical axis of each plot, and those obtained from tract-probability maps are plotted along the horizontal axis. Each plot includes the equality line along which the points would fall if the two methods gave identical results. Overall R^2 values from the mixed-effects model in equation (1), shown in the bottom-right corner of each plot, are a measure of the degree to which conventional-tractography results can be predicted from the tract-probability-map results. Each row presents results for a different MRI index.

The data in Figure 4 indicate that the relationship between MRI indices derived from the two methods is not straightforward, nor is it completely explained by the degree of overlap between participant-specific tract-probability maps and the atlas. Instead, this relationship depends on the specific MRI index and tract under consideration and, to a lesser extent, the presence or absence of MS. Thus, if conventional-tractography results are viewed as a reference standard, FA derived from the tract-probability-map method is biased downward for the corpus callosum and corticospinal tract (the points fall mostly above the diagonal); upward for the optic tract; and downward or upward, depending on the MRI-index value, for the optic radiation. Similar patterns are observed for mean and perpendicular diffusivity. Parallel diffusivity and MTR are the least biased MRI indices, and qT2 is biased upward for all tracts except for the optic tract.

Overall, the strongest association between MRI indices derived from the two methods was for MTR and FA in the corpus callosum and optic radiation ($R^2 > 0.8$). Comparing R^2 values across

tracts, the tract-probability-map method most accurately predicted conventional-tractography results in these two tracts. In the optic tract, however, there was a very poor correlation between the two methods with respect to diffusion indices (FA, MD, $\lambda_{\parallel}$, and $\lambda_{\perp}$). Results for the corticospinal tract were intermediate, with reasonable agreement between the two methods for MD and MTR. Comparing across different MRI indices, the best agreement between the two methods was consistently observed for MTR and the poorest agreement for $\lambda_{\parallel}$ and qT2.

Figure 5 shows spatially normalized profiles for FA, MD, and MTR along the left corticospinal tract for an individual with relapsing-remitting MS who was scanned six times over a three-year period. The horizontal axis is in normalized units and corresponds to the distance between the pons (left side of each panel) and cortex (right side). Results from the six scans (colored lines) are superimposed on the mean and 95% confidence limits from 27 healthy controls (black line and shaded-gray region, respectively). For each profile, the horizontal line shows the location and spatial extent of an MS lesion in the corona radiata that intersected the corticospinal tract and that caused right hemiparesis at the time of baseline scanning (when there was also enhancement on an associated gadolinium-enhanced T1-weighted scan, not shown; no further enhancement within the lesion occurred during the study period).

The data in Figure 5 illustrate several features of the two tract-reconstruction methods that may affect accuracy and reproducibility. Because each scan is separately analyzed in the conventional tractography method, the apparent location of the lesion along the tract is subject to jitter; for the tract-probability-map method, in which all scans are registered to an atlas, the lesion is located at the same relative position in every profile. However, MD at the lesion site is far more variable in the results from the tract-probability-map method, where it appears to increase over time despite a reduction in the spatial extent of the lesion, a relatively stable MTR for the last 3 scans (magnetization-transfer data were not available for the first 3 scans), and a near-complete clinical recovery that occurred during the first year. In the results from the conventional tractography method, on the other hand, there is little change in MD over the three-year period, suggesting that the diffusion characteristics of the lesion did not truly worsen over time. Outside the lesion, the overall variability in MRI-index values (i.e., across the vertical axis) is clearly lower for the tract-probability-map results.

Figure 6 shows average spatially normalized fractional-anisotropy profiles in the MS group for the two tract-reconstruction methods (red for conventional tractography, black for the tract-probability method). These results recapitulate the data of Figure 4 with respect to overall differences between tract-specific MRI indices for the two methods but indicate that the differences are not uniform across the tract. For example, in the anterior portion of the optic radiation (approximately slices 105–112), and in the internal-capsule and corona-radiata portions of the corticospinal tract (slices 35–50), the two methods give similar results despite pronounced differences elsewhere along the tracts.

This overall difference in variability between the two methods is captured in Figure 7, which plots the average within-participant coefficient of variation for the 14 healthy volunteers who contributed more than one scan (32 total scans; median number of days between baseline and follow-up scan: 294). Most of the points fall above the diagonal, indicating greater variability in the conventional-tractography results for all MRI indices except for qT2 in the corpus callosum, optic tract, and corticospinal tract. Other than for qT2 in the corpus callosum and corticospinal tract and all optic-tract MRI indices (blue points), in which the coefficient of variation is on the order of 2–9%, variability is relatively low, typically 0.7–1.5% for the tract-probability-map method and 0.7–3% for the conventional tractography method. Note that this variability combines the effects of noise in the images with errors in coregistration and tract reconstruction.

The associations between tract-specific MRI indices and disability measures are shown in Figure 8, which plots Pearson correlation coefficients calculated via conventional tractography (vertical axis) and tract-probability maps (horizontal axis). Outside the optic tract, the two methods yield similar results, whether the correlations are moderate (corpus callosum vs. PASAT or optic radiation vs. low-contrast visual acuity) or weak (corticospinal tract vs. 25-foot walking time). In the optic tract, the moderate correlations between MRI indices and visual acuity that are measured with conventional tractography are generally weakened when measured using the tract-probability-map method.

Discussion

In this paper, we offer a comparison between the values of quantitative, tract-specific MRI indices derived from conventional tractography and those derived from an automated, atlas-based, tract-probability-map method. To ascertain the applicability of the automated method in the presence of disease, we included data from both healthy individuals and people with MS.

Of the four tracts we studied, the optic tract appears to be very poorly suited to interrogation by the tract probability method described here, whereas the optic radiation and corpus callosum seem best suited. Major potential factors governing the success of the tract-probability-map method include the quality of registration between the tract in each test subject and the atlas and the relative importance of partial-volume averaging between the tract and adjacent tissue. We have assessed registration quality by the overlap index, which was higher in healthy volunteers than in participants with MS for all tracts (indicating better registration in the healthy volunteers). However, we did not find a straightforward relationship between registration quality and concordance between MRI indices derived from the two methods. For example, although overlap indices were similar for the corpus callosum and optic tract, concordance was much better in the corpus callosum.

For these reasons, we suspect that partial-volume averaging played the major role in determining the success of the tract-probability-map method. This is particularly important in the optic tract, which is small and partially situated within the cerebrospinal fluid of the suprasellar cistern, so that a large number of its component voxels border the cerebrospinal fluid and are therefore subject to partial-volume averaging with water rather than brain tissue. The extent to which such mixed or non-tract voxels are included in the analysis may change from scan to scan, leading to a great deal of scatter in tract-specific MRI indices (second column of Figure 4) and high scan-to-scan variability (blue points in Figure 7). This can partially explain the poor correlation between low-contrast visual acuity and optic-tract-specific MRI indices, particularly when measured using the tract-probability-map method (Figure 8, bottom-right panel).

One possible explanation for the greater success of the tract-probability-map method with the other tracts, especially the optic radiation and corpus callosum, may be their particular size and morphology. These tracts have a sheet-like configuration and are less tubular than both the optic tract and corticospinal tract (Mori et al., 2005). Because of this sheet-like configuration, registration errors in some directions (i.e., within the plane of the tract) may have less of an impact on tract-specific MRI indices. Moreover, the adverse impact of partial-volume averaging may be less relevant in these tracts if the MRI properties of adjacent tissue are similar; in the case of the optic tract, the adjacent tissue is cerebrospinal fluid, the MRI properties of which are quite different from those of brain parenchyma.

The degree to which MRI indices derived from tract-probability maps can predict the corresponding indices derived from conventional tractography differs from index to index

(Figure 4). Because MTR maps of the white matter are relatively homogeneous, MTR indices derived from the two methods are highly correlated, with values of R^2 reaching 0.83: they are less sensitive to inadvertent inclusion of voxels from outside a given tract. The relatively poor correlation seen for qT2 is due, at least in part, to poor coregistration between these images, which are generated using a fast-spin-echo readout scheme, and the more distorted echo-planar images used for the DTI and MTR scans. These registration errors make it more likely that cerebrospinal-fluid voxels with high T2 values are included in the qT2 analysis, since the masking of those voxels is based on their positions in the DTI scans.

Overall scan-to-scan variability in MRI indices, measured by the coefficient of variation across repeated scans in healthy volunteers, is acceptable for both methods but is substantially better for indices derived from tract-probability maps (generally 0.7–1.5% vs. 0.7–3% for conventional tractography). This is demonstrated graphically in Figure 7, in which most points fall above the main diagonal. The variability is lower when tractography is not performed separately in each case because approximately the same amount of tissue is included in all scans (except for cerebrospinal-fluid voxels, which are removed on a case-by-case basis). The exception to this rule is qT2 (in all tracts except for the optic radiation), for which variability is greater in the tract-probability-map results due to poorer registration, as discussed above.

Both methods have drawbacks. For conventional tractography to successfully reconstruct a tract, the diffusion characteristics (i.e., anisotropy and directionality) need to be relatively stable at every point. A single voxel with low FA or an eigenvector that is too highly rotated will cause premature truncation of the tract. This is a particular problem in the presence of disease, where tracts can sometimes fail to be reconstructed at lesion sites (Lin et al., 2005; Reich et al., 2007). This problem is avoided in the atlas-based, tract-probability-map approach, in which tractography is not performed separately for each scan. However, the tract-probability-map method is itself limited by imperfect registration to the atlas, which is a particular problem in diseases that cause brain atrophy or regions of abnormal signal intensity (e.g., MS lesions). This results in inclusion of tissue from outside the tract of interest. By setting an FA threshold below which voxels (usually those containing cerebrospinal fluid and gray matter; here, voxels with FA<0.25) are excluded from analysis, we have partially mitigated this problem. However, this procedure does not limit the inadvertent inclusion of tissue from adjacent white matter tracts with higher FA values, as might have occurred in the example shown in Figure 5. Moreover, applying this threshold may still spuriously remove some of the most damaged voxels within a tract, in which FA is typically lowest, although this was not a major problem in our data.

These considerations partially account for the variable and imperfect relationship between MRI indices derived from the two methods. It is therefore somewhat surprising that the methods yield similar correlation coefficients for the relationship between MRI indices and clinical data (Figure 8). Although these correlation coefficients are only moderately strong (on the order of 0.4 or less), the results suggest that the specific method for obtaining tract-specific MRI indices matters relatively little outside tracts for which the tract-probability-map method fails entirely, such as the optic tract. This is encouraging, because it suggests that the fully automated tract-probability-map method can be used in studies that seek to examine the relationship between clinical disability and the imaging properties of specific functional systems within the white matter. As a corollary, the results also suggest that the exclusion from the analyses of even the most damaged portions of tracts, as might occur with conventional tractography in the presence of lesions, may not have a profound impact on the results of such studies.

The automated tract-probability-map method presented would certainly be improved by the use of registration techniques with more degrees of freedom, including nonlinear techniques, both for creation of the tract-probability maps and for analysis of test scans. Within-participant

registration would probably have improved with affine rather than rigid-body registration, particularly for registration of the fast-spin-echo T2-weighted scans to the echo-planar diffusion-weighted scans, but also for registration of the diffusion-weighted scans to the b_0 scan due to eddy-current-related distortions (although the use of parallel imaging limited the extent of such distortions in our data). High degree-of-freedom and nonlinear techniques could potentially improve the extent to which the tract-probability maps correspond to the true tracts in each test scan at the expense of additional processing time. However, the elasticity of a nonlinear transformation between images from diseased and healthy brains could yield spurious registration results, and our initial experience with these techniques indicates that this is indeed an important factor to consider. Moreover, given the already high degree of correspondence between the correlation coefficients linking disability scores to tract-specific MRI indices, improved registration techniques may not ultimately matter a great deal in this context. Nevertheless, we are currently developing and testing a nonlinear registration approach in our lab.

Our study fits well with previous reports describing the use of tract-probability maps in disease. Pagani et al. (Pagani et al., 2005) studied the corticospinal tract using a linear-registration approach for creating the tract-probability maps and applying them to MS scans. They did not directly compare the results of the two tract-reconstruction methods but demonstrated that diffusion indices in the corticospinal tract differ between healthy and MS brains even when using tract-probability maps. One potentially important difference between their approach and ours is that they transformed all scans to the Montreal Neurological Institute T1-weighted atlas, whereas we selected a single reference scan from a healthy volunteer. Because our test and reference scans had similar echo-planar-imaging-related distortions, this selection should reduce the effect of coregistration errors. In related work (Pagani et al., 2007b), the same group demonstrated the potential of a nonlinear extension of the method to assess tract-specific volume changes, something we did not attempt in the work described here.

Previous work from our own group (Hua et al., 2008) also used a standard atlas space and compared the results of conventional tractography and tract-probability mapping for a group of healthy volunteers and a single individual with MS. The current paper represents an extension of those results with additional methodological optimization to improve the method's performance in the face of disease (particularly spurious registration results resulting from brain atrophy). Indeed, the tract profiles shown in Figure 5 of this paper are from the same individual depicted in Figure 7 of Hua et al. As mentioned above, we are in the process of merging this previous work with a highly nonlinear registration method called Large Deformation Diffeomorphic Metric Mapping, the application of which to regional brain mapping has been recently described (Oishi et al., 2009).

Additional improvements on and refinements of the tract-probability-map method were recently described and applied to scans from individuals with temporal lobe epilepsy (Hagler et al., 2009). The approach described there involves linear registration of the DTI scans to the T1-weighted scans for each participant followed by nonlinear registration of the T1-weighted scans to a standard T1-space. Additional modifications include intensity corrections and fiber-orientation regularization. Consistent with the results reported in this paper, tract-specific FA values derived from conventional tractography and tract-probability mapping were found to be quite similar to one another, and the reproducibility of FA values was improved in the tract-probability-map approach.

Alternative approaches to deriving tract-specific MRI indices that do not use tractography or tract-probability mapping have been described and are likely to provide complementary results to those presented here. A prominent example is tract-based spatial statistics (Smith et al., 2006). This method considers all the data in a given study to derive a tract-skeleton map, and

individual voxels are then assigned to particular portions of that skeleton. Intersession-variability results are similar to those reported here. More recently, our collaborators have introduced a clustering method for DTI data that can segregate the brain into individual tracts using prior information and local similarity measures (Bazin et al., 2009). In preliminary analysis, this method appears to work well even in the presence of MS lesions and may represent an especially promising approach.

In summary, we have described and evaluated the performance of a tract-probability-map method that, for selected tracts, yields tract-specific MRI indices that are moderately to strongly correlated with those of conventional tractography. This method has the advantage of being fully automated and not subject to tractography failures that may occur in the presence of lesions. Certain tracts, especially the corpus callosum and optic radiation, appear to be most accurately realized by this method. Via registration of non-DTI scans to the DTI-based tract atlas, non-DTI indices (here, MTR and qT2) can also be assessed in a tract-specific fashion. Results from the tract-probability-map method are less variable than those derived from conventional tractography but, importantly, are similarly correlated to disability scores in MS.

Acknowledgments

We thank Terri Brawner, Brian Caffo, Hormuzdiyar Dasenbrock, Jonathan Farrell, Sheena Farrell, Eliza Gordon-Lipkin, Daniel Harrison, Kegang Hua, Kathleen Kahl, Ivana Kusevic, Bennett Landman, and Seth Smith for assistance with data collection and analysis.

Funding sources: NIH grants K99NS064098, P41RR015241, R01AG020012, and K01EB009120; and National Multiple Sclerosis Society grant TR3760A3. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

References

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

Bazin, P-L.; Bogovic, J.; Reich, DS.; Pham, DL.; Prince, J. Belief propagation based segmentation of white matter tracts in DTI; International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI); London, England. 2009.

Burgel U, Madler B, Honey CR, Thron A, Gilsbach J, Coenen VA. Fiber tracking with distinct software tools results in a clear diversity in anatomical fiber tract portrayal. Cen Eur Neurosurg 2009;70:27–35. [PubMed: 19191204]

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 U S A 1999;96:10422–10427. [PubMed: 10468624]

Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, Syndulko K, Weinshenker BG, Antel JP, Confavreux C, Ellison GW, Lublin F, Miller AE, Rao SM, Reingold S, Thompson A, Willoughby E. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain 1999;122(Pt 5):871–882. [PubMed: 10355672]

Dasenbrock HH, Smith SA, Ozturk A, Farrell SK, Calabresi PA, Reich DS. Diffusion tensor imaging of the optic tract in multiple sclerosis: association with retinal thinning and visual disability. J Neuroimaging submitted. 2009

Goodin DS. Magnetic resonance imaging as a surrogate outcome measure of disability in multiple sclerosis: have we been overly harsh in our assessment? Ann Neurol 2006;59:597–605. [PubMed: 16566022]

Hagler DJ, Ahmadi ME, Kuperman J, Holland D, McDonald CR, Halgren E, Dale AM. Automated white matter tractography using a probabilistic diffusion tensor atlas: Application to temporal lobe epilepsy. Hum Brain Mapp 2009;30:1535–1547. [PubMed: 18671230]

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]

Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, Calabresi PA, Pekar JJ, van Zijl PC, Mori S. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. Neuroimage 2008;39:336–347. [PubMed: 17931890]

Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 2002;17:825–841. [PubMed: 12377157]

Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–1452. [PubMed: 6685237]

Landman BA, Farrell JA, Jones CK, Smith SA, Prince JL, Mori S. Effects of diffusion weighting schemes on the reproducibility of DTI-derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5T. Neuroimage 2007;36:1123–1138. [PubMed: 17532649]

Lee JS, Han MK, Kim SH, Kwon OK, Kim JH. Fiber tracking by diffusion tensor imaging in corticospinal tract stroke: Topographical correlation with clinical symptoms. Neuroimage 2005;26:771–776. [PubMed: 15955486]

Lin X, Tench CR, Morgan PS, Constantinescu CS. Use of combined conventional and quantitative MRI to quantify pathology related to cognitive impairment in multiple sclerosis. J Neurol Neurosurg Psychiatry 2008;79:437–441. [PubMed: 17673493]

Lin X, Tench CR, Morgan PS, Niepel G, Constantinescu CS. 'Importance sampling' in MS: use of diffusion tensor tractography to quantify pathology related to specific impairment. J Neurol Sci 2005;237:13–19. [PubMed: 16109428]

Mori S, Crain BJ, Chacko VP, van Zijl PC. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann Neurol 1999;45:265–269. [PubMed: 9989633]

Mori S, Oishi K, Jiang H, Jiang L, Li X, Akhter K, Hua K, Faria AV, Mahmood A, Woods R, Toga AW, Pike GB, Neto PR, Evans A, Zhang J, Huang H, Miller MI, van Zijl P, Mazziotta J. Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. Neuroimage 2008;40:570. [PubMed: 18255316]

Mori, S.; Wakana, S.; van Zijl, PCM.; Nagae-Poetscher, LM. MRI Atlas of Human White Matter. Amsterdam: Elsevier; 2005.

Oishi K, Faria A, Jiang H, Li X, Akhter K, Zhang J, Hsu JT, Miller MI, van Zijl PC, Albert M, Lyketsos CG, Woods R, Toga AW, Pike GB, Rosa-Neto P, Evans A, Mazziotta J, Mori S. Atlas-based whole brain white matter analysis using large deformation diffeomorphic metric mapping: application to normal elderly and Alzheimer's disease participants. Neuroimage 2009;46:486–499. [PubMed: 19385016]

Ozturk A, Smith SA, Gordon-Lipkin E, Harrison DM, Shiee N, Pham DL, Caffo BS, Calabresi PA, Reich DS. MRI of the corpus callosum in multiple sclerosis: association with disability. Multiple Sclerosis. 2009 in press.

Pagani E, Bammer R, Horsfield MA, Rovaris M, Gass A, Ciccarelli O, Filippi M. Diffusion MR imaging in multiple sclerosis: technical aspects and challenges. AJNR Am J Neuroradiol 2007a;28:411–420. [PubMed: 17353305]

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]

Pagani E, Horsfield MA, Rocca MA, Filippi M. Assessing atrophy of the major white matter fiber bundles of the brain from diffusion tensor MRI data. Magn Reson Med 2007b;58:527–534. [PubMed: 17763353]

Parker GJ, Haroon HA, Wheeler-Kingshott CA. A framework for a streamline-based probabilistic index of connectivity (PICo) using a structural interpretation of MRI diffusion measurements. J Magn Reson Imaging 2003;18:242–254. [PubMed: 12884338]

Reich DS, Smith SA, Gordon-Lipkin EM, Ozturk A, Caffo BS, Balcer LJ, Calabresi PA. Damage to the optic radiation in multiple sclerosis is associated with retinal injury and visual disability. Arch Neurol 2009;66:998–1006. [PubMed: 19667222]

Reich DS, Smith SA, Jones CK, Zackowski KM, van Zijl PC, Calabresi PA, Mori S. Quantitative characterization of the corticospinal tract at 3T. AJNR Am J Neuroradiol 2006;27:2168–2178. [PubMed: 17110689]

Reich DS, Smith SA, Zackowski KM, Gordon-Lipkin EM, Jones CK, Farrell JA, Mori S, van Zijl PC, Calabresi PA. Multiparametric magnetic resonance imaging analysis of the corticospinal tract in multiple sclerosis. Neuroimage 2007;38:271–279. [PubMed: 17870615]

Reich DS, Zackowski KM, Gordon-Lipkin EM, Smith SA, Chodkowski BA, Cutter GR, Calabresi PA. Corticospinal tract abnormalities are associated with weakness in multiple sclerosis. AJNR Am J Neuroradiol 2008;29:333–339. [PubMed: 17974617]

Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17:143–155. [PubMed: 12391568]

Smith SM, Jenkinson M, Johansen-Berg H, Rueckert D, Nichols TE, Mackay CE, Watkins KE, Ciccarelli O, Cader MZ, Matthews PM, Behrens TE. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 2006;31:1487–1505. [PubMed: 16624579]

Wakana S, Caprihan A, Panzenboeck MM, Fallon JH, Perry M, Gollub RL, Hua K, Zhang J, Jiang H, Dubey P, Blitz A, van Zijl P, Mori S. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 2007;36:630–644. [PubMed: 17481925]

Wheeler-Kingshott CA, Cercignani M. About "axial" and "radial" diffusivities. Magn Reson Med 2009;61:1255–1260. [PubMed: 19253405]

Witelson SF. Hand and sex differences in the isthmus and genu of the human corpus callosum. A postmortem morphological study. Brain 1989;112(Pt 3):799–835. [PubMed: 2731030]

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]

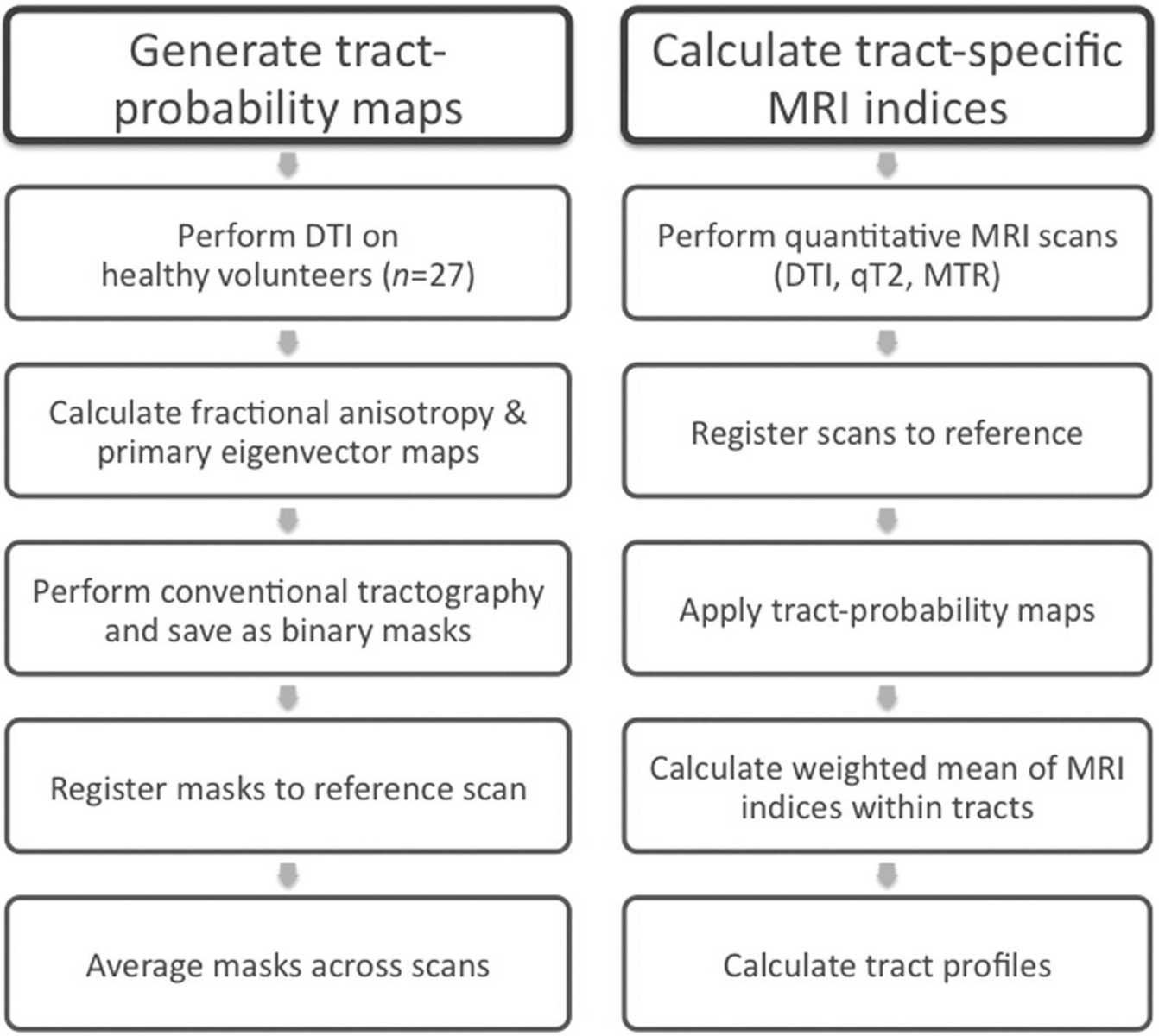

Figure 1.
Diagram of the process for creating tract-probability maps (left column) and applying them to data from test subjects (right column).

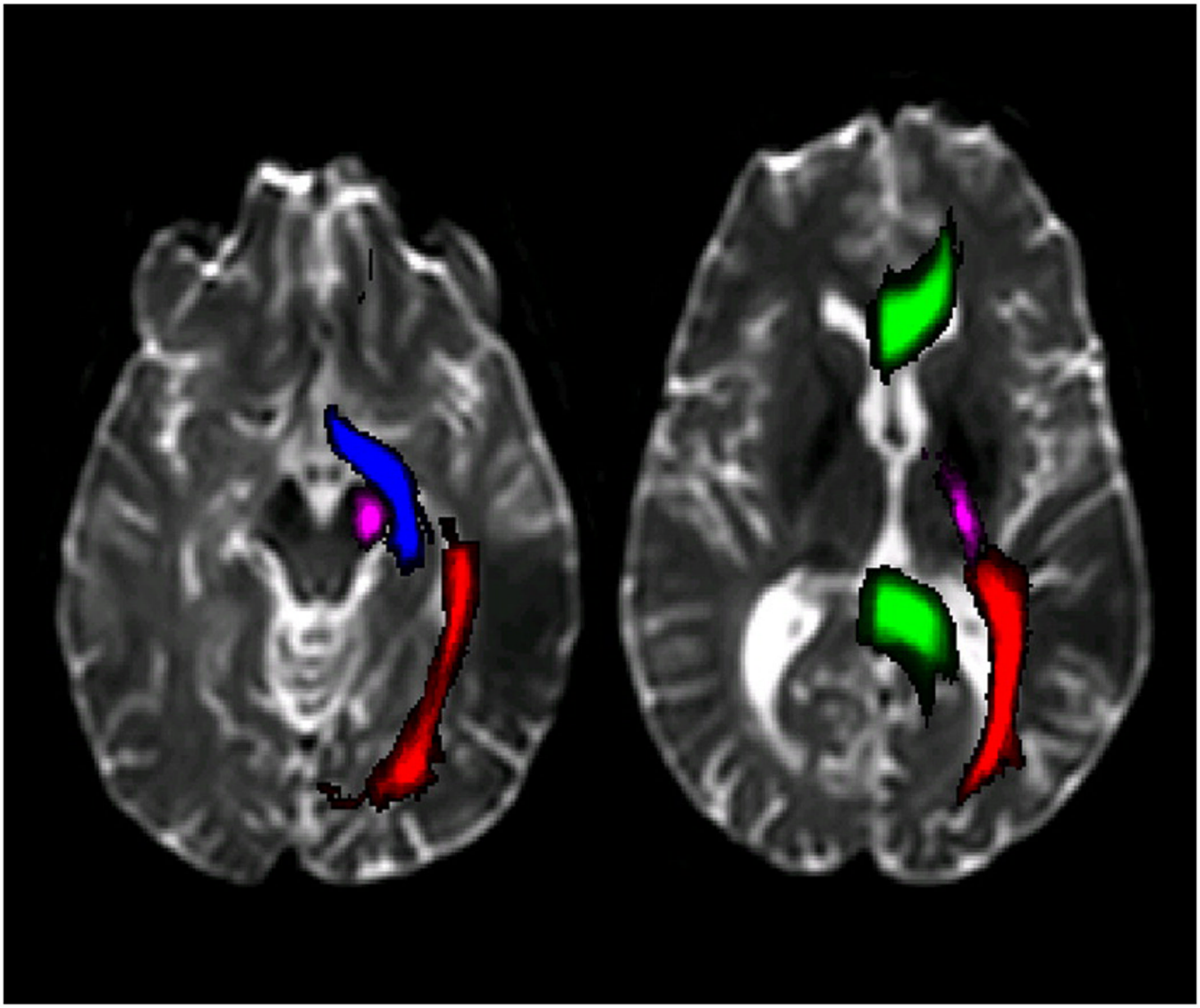

Figure 2.
Representative axial sections from the minimally diffusion-weighted (b_0) reference scan (35-year-old healthy woman) with superimposed tract-probability maps overlaid on the left side of the brain. *Red*, optic radiation; *green*, corpus callosum; *blue*, optic tract; *purple*, corticospinal tract. The color intensity corresponds to the probability that the underlying voxel is within the tract of interest.

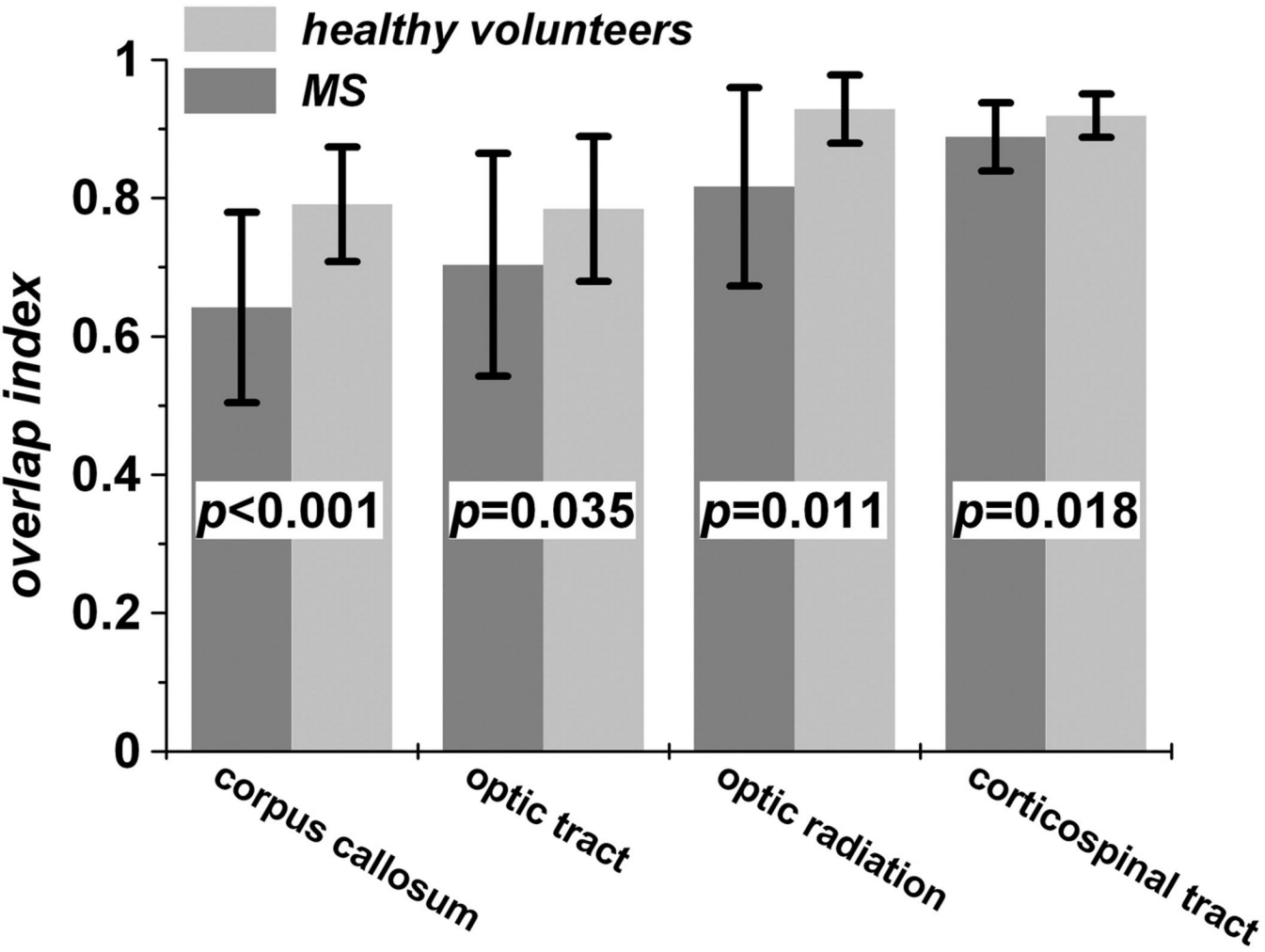

Figure 3.
For the tract-probability-map method, overlap indices (mean ± standard deviation) for each tract comparing healthy volunteers to individuals with multiple sclerosis. An overlap index of 1 indicates perfect correspondence between the reconstructed tract and the atlas, whereas an overlap index of 0 indicates no correspondence. The *p*-values account for differences in age between the two groups and for multiple scans per participant.

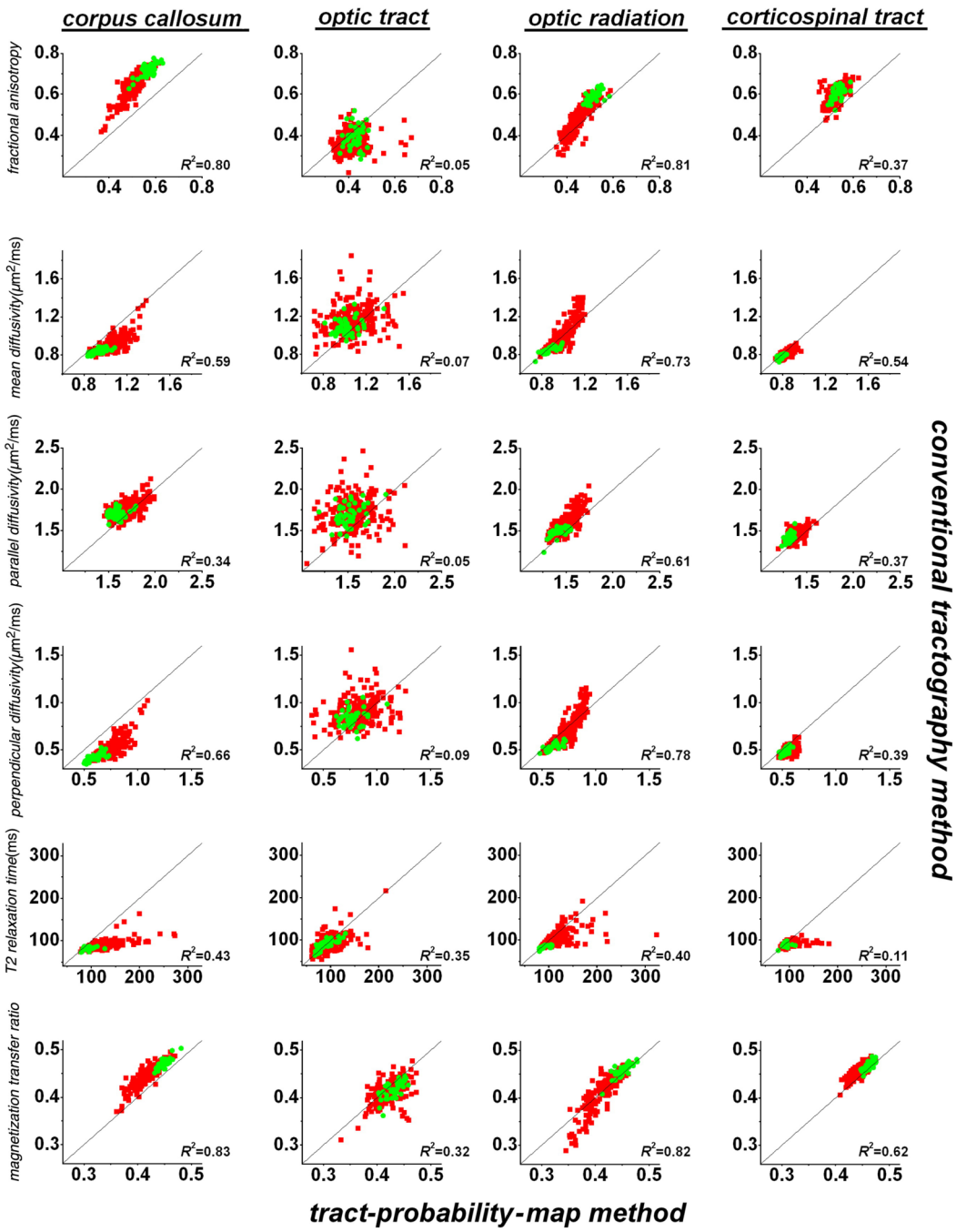

Figure 4.

MRI indices derived from conventional tractography (vertical axes) compared to those derived from the tract-probability-map method (horizontal axes). Each row depicts a different MRI index and each column a different tract. *Red points*, multiple-sclerosis scans. *Green points*, healthy-volunteer scans. The mixed-effects-model values of R^2 are shown in each panel.

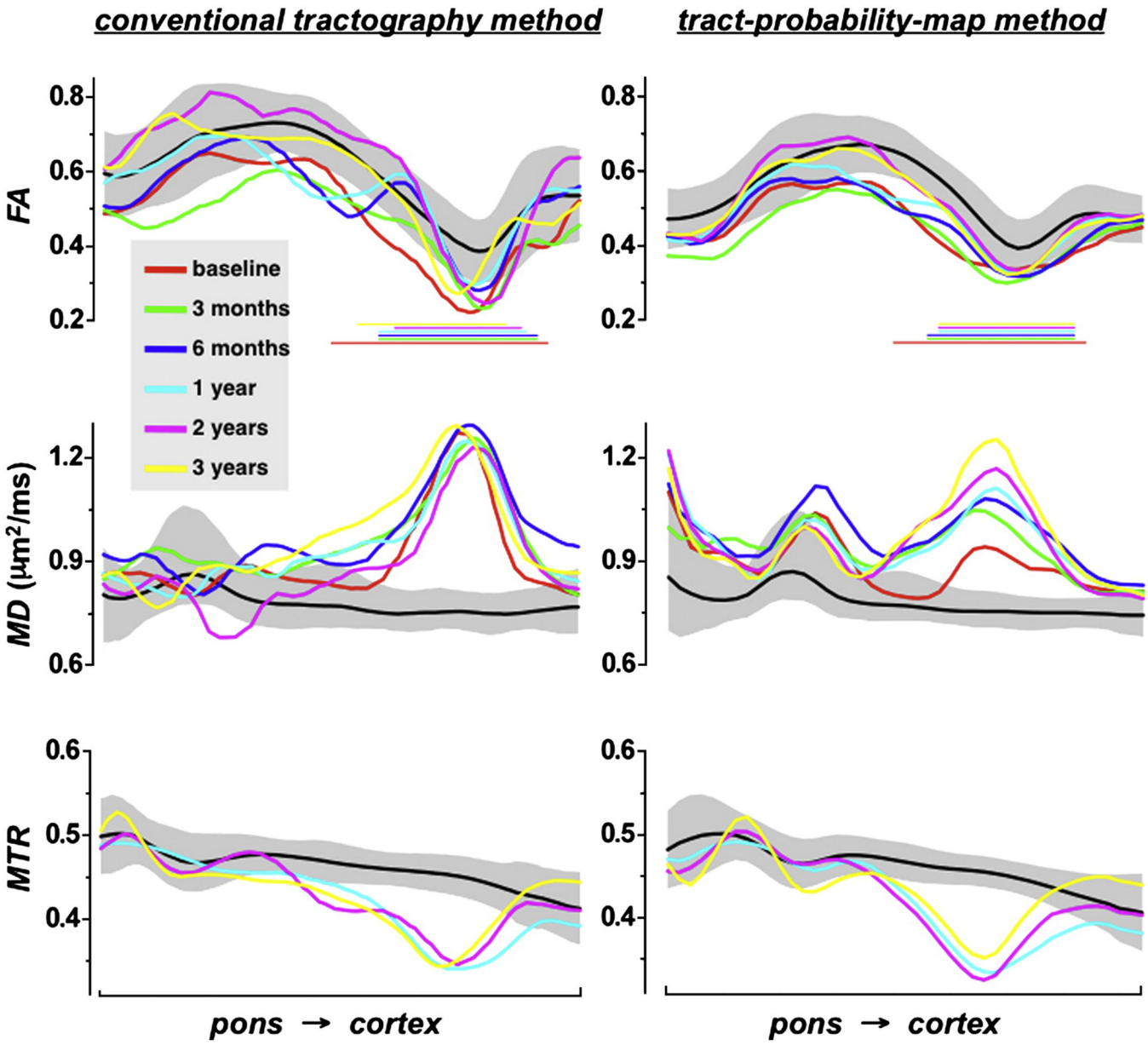

Figure 5.

Spatially normalized tract profiles for fractional anisotropy (FA; top row) and mean diffusivity (MD; middle row), and magnetization transfer ratio (MTR; bottom row) of the left corticospinal tract from a man with relapsing-remitting multiple sclerosis. The man was 33 years old when first scanned, and 5 follow-up scans were obtained over a period of 3 years. The horizontal lines denote the position of an MS lesion that was acute at the time of first scanning and that caused temporary right hemiparesis. Tract profiles derived from conventional tractography are shown in the left column and from the tract-probability-map method in the right column. The horizontal axis depicts normalized position along the tract, ranging from the pons on the left to the subcortical white matter on the right. The solid black line shows the mean tract profile from healthy volunteers and the gray zone corresponds to the 95% confidence limits of the mean.

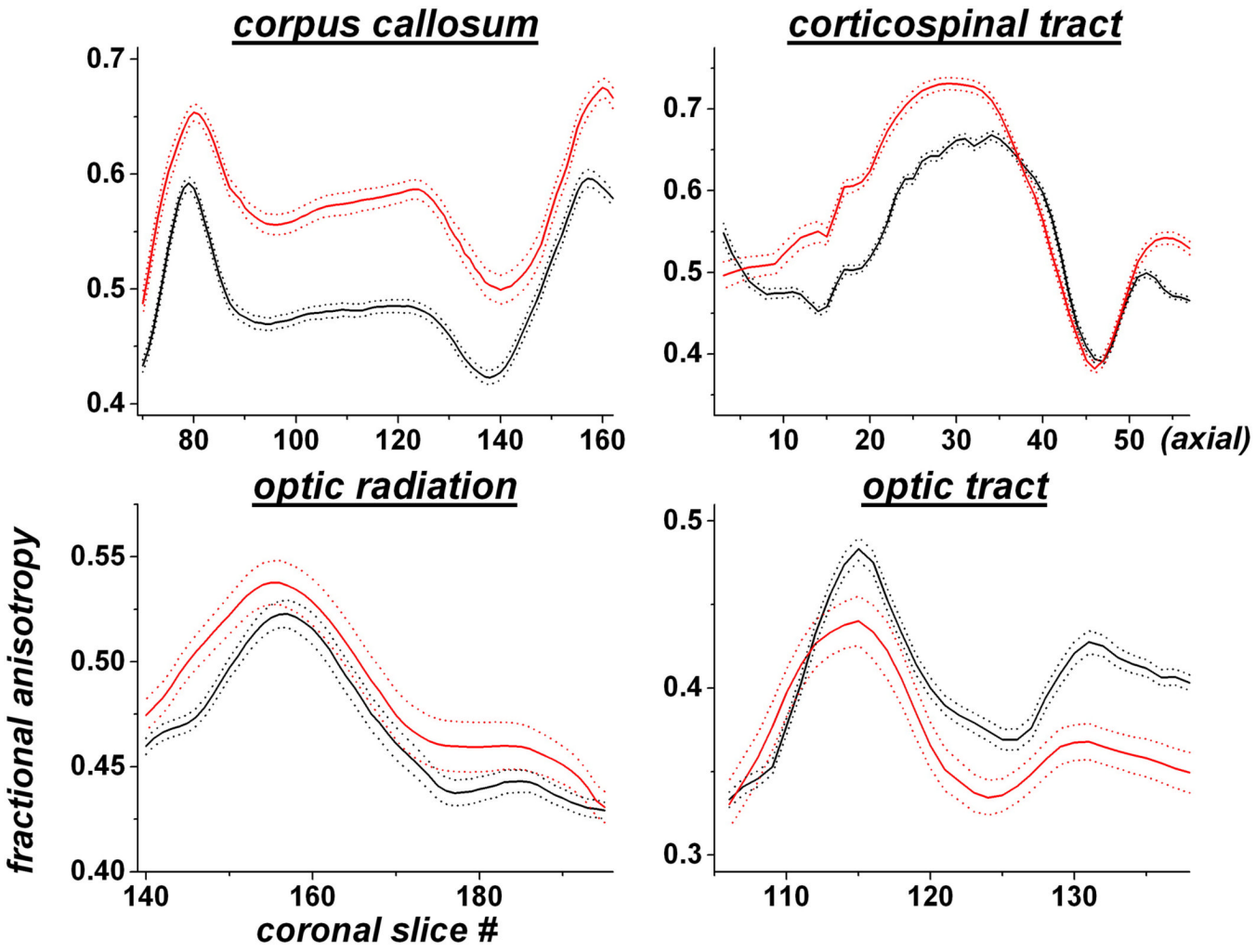

Figure 6.

Spatially normalized tract profiles (mean ± 2 standard errors) for fractional anisotropy across the multiple-sclerosis cohort. *Red*, conventional tractography; *black*, tract-probability-map method. The horizontal axis reflects the slice number in the atlas space (coronal for the corpus callosum, optic tract, and optic radiation; axial for the corticospinal tract).

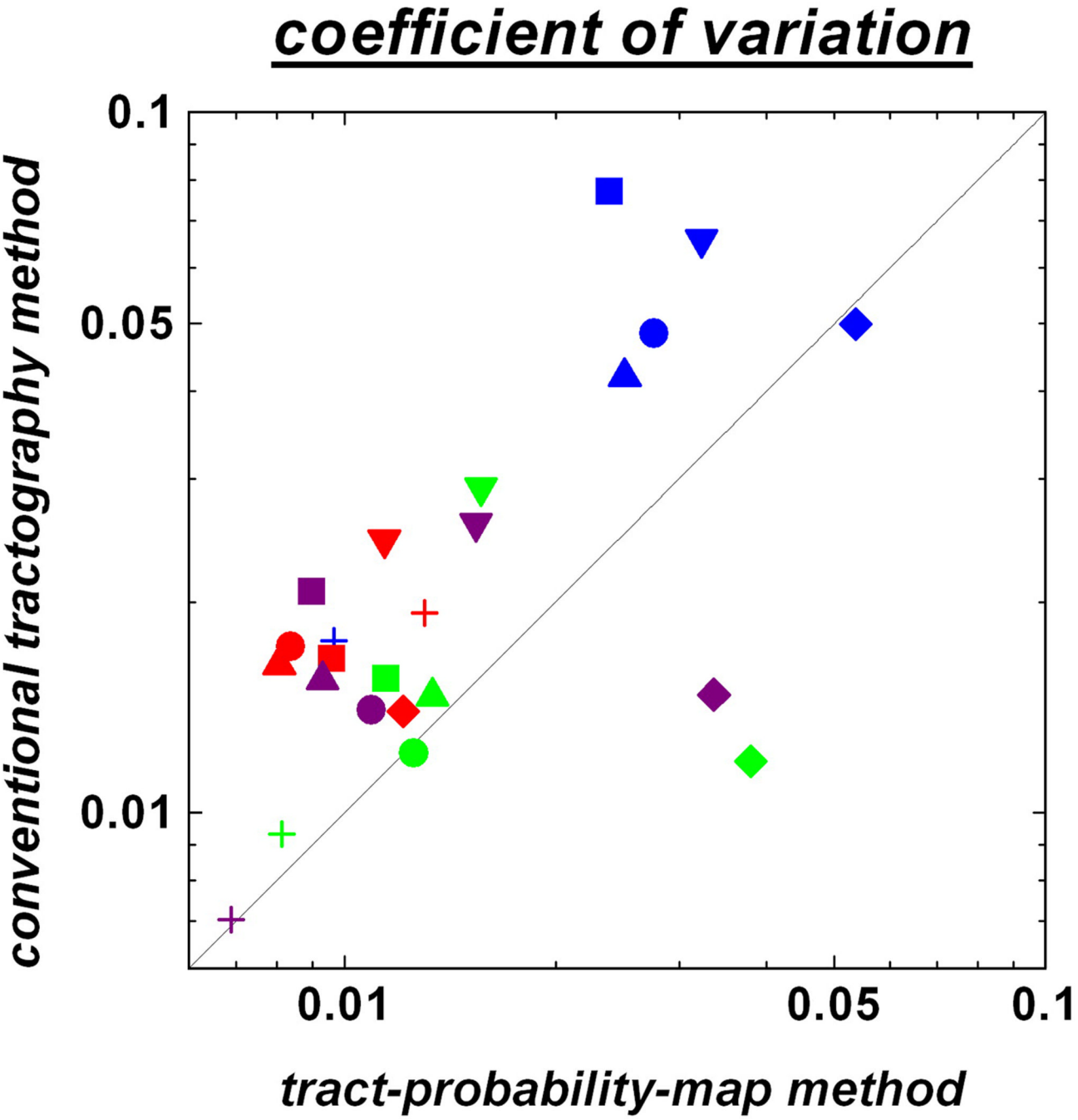

Figure 7.
Coefficient of variation of MRI indices derived from conventional tractography (vertical axis) and tract-probability maps (horizontal axis) across repeated scans of the same individual. Data are from 32 scans of 14 healthy volunteers, and the number of scans ranges from 2 to 4. Tracts are coded by color: *Red*, optic radiation; *green*, corpus callosum; *blue*, optic tract; *purple*, corticospinal tract. MRI indices are coded by symbol type: fractional anisotropy (■); mean diffusivity (●); parallel diffusivity (▲); perpendicular diffusivity (▼); quantitative T2 (♦); magnetization transfer ratio (+).

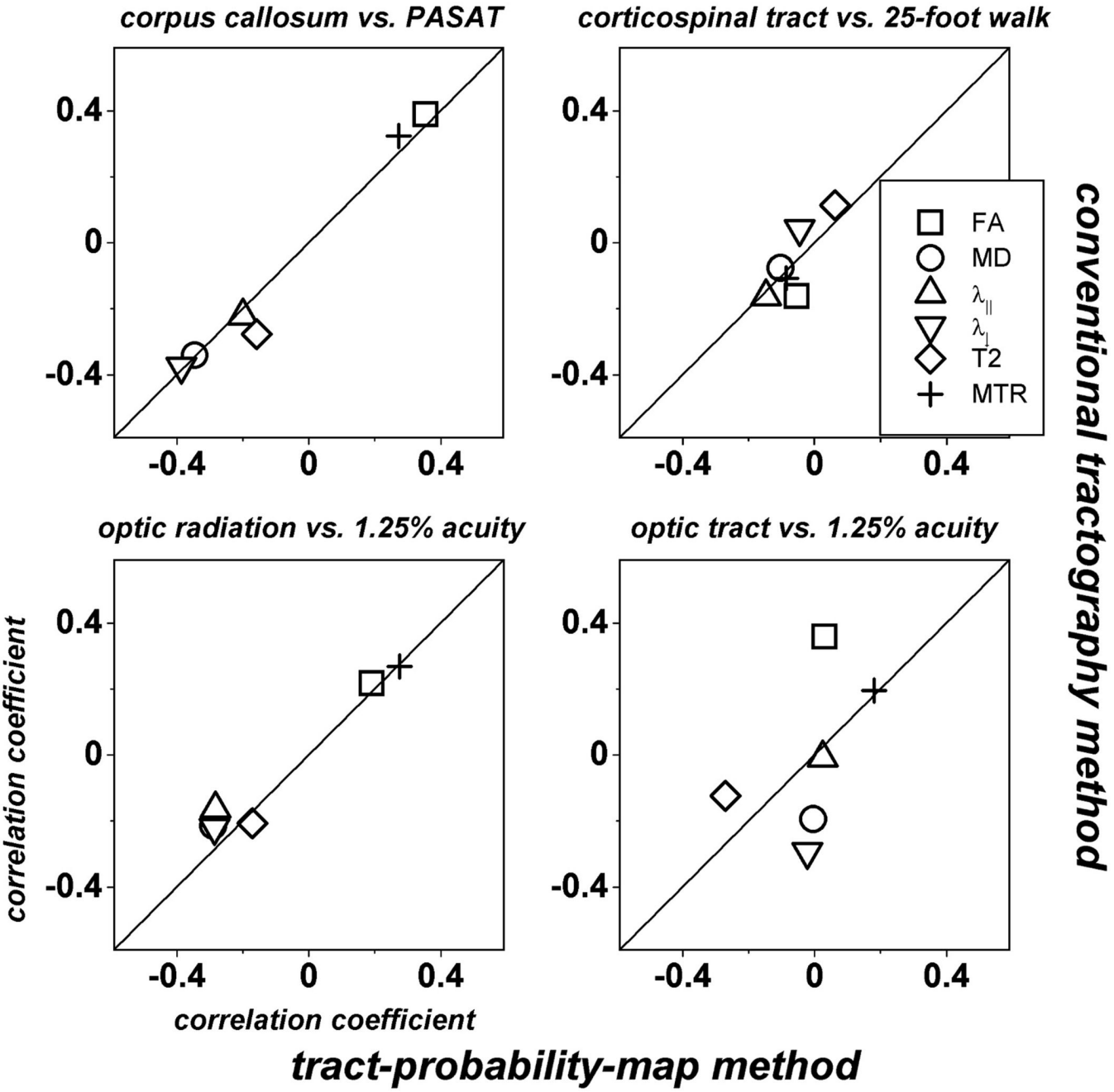

Figure 8.

Correlation coefficients relating tract-specific MRI indices derived from conventional tractography (vertical axes) and the tract-probability-map method (horizontal axes) to clinical disability scores appropriate to that tract. Note that the correlation coefficients derived from the two methods are similar even when the coefficients themselves are small (e.g., corticospinal tract vs. 25-foot walk), except for the case of the optic tract, where the tract-probability-map method fails.