Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited

Cibu Thomas[a,b,1], Frank Q. Ye[c,d], M. Okan Irfanoglu[a,b], Pooja Modi[a], Kadharbatcha S. Saleem[e], David A. Leopold[c,d], and Carlo Pierpaoli[a,b]

[a]Program on Pediatric Imaging and Tissue Sciences, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, 20892; [b]Center for Neuroscience and Regenerative Medicine, Uniformed Services University of the Health Sciences, Bethesda, MD 20814; [c]Neurophysiology Imaging Facility, National Institute of Mental Health, National Institute of Neurological Disorders and Stroke, and National Eye Institute, Bethesda, MD 20892; [d]Section on Cognitive Neurophysiology and Imaging and [e]Section on Cognitive Neuroscience, Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, MD, 20892

Edited* by Leslie G. Ungerleider, National Institute of Mental Health, Bethesda, MD, and approved September 22, 2014 (received for review March 26, 2014)

Tractography based on diffusion-weighted MRI (DWI) is widely used for mapping the structural connections of the human brain. Its accuracy is known to be limited by technical factors affecting in vivo data acquisition, such as noise, artifacts, and data undersampling resulting from scan time constraints. It generally is assumed that improvements in data quality and implementation of sophisticated tractography methods will lead to increasingly accurate maps of human anatomical connections. However, assessing the anatomical accuracy of DWI tractography is difficult because of the lack of independent knowledge of the true anatomical connections in humans. Here we investigate the future prospects of DWI-based connectional imaging by applying advanced tractography methods to an ex vivo DWI dataset of the macaque brain. The results of different tractography methods were compared with maps of known axonal projections from previous tracer studies in the macaque. Despite the exceptional quality of the DWI data, none of the methods demonstrated high anatomical accuracy. The methods that showed the highest sensitivity showed the lowest specificity, and vice versa. Additionally, anatomical accuracy was highly dependent upon parameters of the tractography algorithm, with different optimal values for mapping different pathways. These results suggest that there is an inherent limitation in determining long-range anatomical projections based on voxel-averaged estimates of local fiber orientation obtained from DWI data that is unlikely to be overcome by improvements in data acquisition and analysis alone.

diffusion MRI | tractography | white matter | tracer | validation

The creation of a comprehensive map of the connectional neuroanatomy of the human brain would be a fundamental achievement in neuroscience. However, despite the numerous efforts to date (for a historical review, see ref. 1), creating this map remains a challenge. A major limitation is that the current gold-standard technique for mapping structural connections, which requires the injection of axonal tracers, cannot be used in humans. The introduction of diffusion-weighted MRI (DWI) (2–4) and the subsequent advent of diffusion tensor MRI (DTI) (5) opened the possibility of exploring the structural properties of white matter in the living human brain (6). Local DWI measures are used clinically for the early detection of stroke and for the characterization of neurological disorders such as multiple sclerosis, epilepsy, and brain gliomas, among others (7). In addition, tractography approaches (8–12) that can infer structural brain connectivity based on brain-wide local DWI measurement have been developed (for reviews, see refs. 13 and 14). The success of DWI tractography as a method for studying fiber trajectories has led to a systematic characterization of large white-matter pathways of the living human brain (e.g., ref. 15), and now it is used routinely to provide a structural explanation for aspects of human brain function (16).

A major limitation of DWI tractography is that its characterization of axonal pathways is based on indirect information and numerous assumptions. Local white matter orientation profiles are based on the statistical displacement profile (i.e., diffusion propagator) of water molecules in brain tissue on the coarse scale of a voxel, and fiber trajectories are inferred based on the adjacency of similar diffusion profiles. This approach differs fundamentally from conventional tract-tracing approaches in animals, which involve the physical transport of traceable molecules through the cells' axoplasm over a large distance. Because these molecules occupy positions within the axon, it sometimes is possible to reconstruct the trajectory of individual neurons through the white matter (e.g., ref. 17). Given the inherent coarseness of DWI tractography, it can be argued that the prospect of using this method to reconstruct complex axonal pathways accurately in the human brain, in a manner similar to that used for molecular tracers in animals, is likely to be intrinsically problematic. Indeed, the limitations of DWI tractography techniques have been noted since their inception (8), and the anatomical accuracy of results from tractography based on the tensor model has been shown to be mixed (18). This inaccuracy has been attributed to two main factors. The first relates to the assumptions underlying tractography algorithms. For example, it has long been recognized that a simple tensor model (19) of local diffusion leads to problems in certain white matter regions where fibers cross within individual voxels. As a remedy, high angular resolution diffusion imaging (HARDI) methods (e.g., refs. 20–24) have been developed to enable better characterization of the diffusion displacement profile and to improve the accuracy of tractography. The second factor limiting accuracy stems from the low quality of clinical DWI data because of various sources of noise. Eddy current distortions,

Significance

Diffusion-weighted MRI (DWI) tractography is widely used to map structural connections of the human brain in vivo and has been adopted by large-scale initiatives such as the human connectome project. Our results indicate that, even with high-quality data, DWI tractography alone is unlikely to provide an anatomically accurate map of the brain connectome. It is crucial to complement tractography results with a combination of histological or neurophysiological methods to map structural connectivity accurately. Our findings, however, do not diminish the importance of diffusion MRI as a noninvasive tool that offers important quantitative measures related to brain tissue microstructure and white matter architecture.

Author contributions: C.T., F.Q.Y., M.O.I., D.A.L., and C.P. designed research; C.T., F.Q.Y., M.O.I., K.S.S., D.A.L., and C.P. performed research; C.T., F.Q.Y., M.O.I., P.M., and K.S.S. contributed new reagents/analytic tools; C.T., M.O.I., K.S.S., and C.P. analyzed data; C.T., F.Q.Y., D.A.L., and C.P. wrote the paper; and C.P. supervised the project.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence should be addressed. Email: cibu.thomas@nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1405672111/-/DCSupplemental.

subject motion, physiological noise (see ref. 25 for a review), and susceptibility artifacts from echo planar imaging (EPI) (26) all lead to poor local characterization of diffusion and, consequently, to incorrect tractography results. Continuing advances in sequence design, MRI gradient hardware, and postprocessing correction schemes have overcome many of the initial problems (27) and have led to the belief that further acquisition improvements will result in more precise mapping of structural connections in the human brain (28). In fact, the assumption underlying many recent initiatives to map structural brain connectivity from DWI data is that improved image data quality and sophisticated diffusion modeling approaches will result in anatomically accurate maps of white matter connections (29). The goal of the present study is to investigate the validity of this assumption.

To achieve this goal, we acquired high angular resolution DWI data from a normal adult rhesus macaque brain, ex vivo, at a spatial resolution of 250 microns (isotropic). This dataset is ideal for exploring the limits of DWI tractography because of its high signal-to-noise ratio (SNR) (for SNR computation, see *SI Materials and Methods*) and the almost complete absence of experimental confounds and artifacts such as those originating from patient motion, noise, cardiac pulsation, and EPI distortion that are typically encountered in in vivo studies. Using the axonal tracer results from a well-known atlas (17) as reference, we measured the sensitivity (i.e., the ability to detect true connections) and specificity (i.e., the ability to avoid false connections) of several DWI tractography implementations representative of the current state of the art. This approach allowed us to investigate whether sophisticated diffusion modeling techniques, when applied to DWI data of exceptional quality, would yield accurate maps of axonal connections.

Results

We selected two independent datasets from a previous study (17) comprising anatomical locations of tracer-labeled ROIs that were produced by anterograde tracer injections placed within (*i*) the precentral gyrus (PCG or area 4) corresponding to the foot region of the motor cortex (i.e., Case 28 in ref. 17, p. 322) and (*ii*) the rostroventral part of the occipital region corresponding to the ventral part of area V4 (V4v) and the adjacent ventral area V3 (i.e., Case 21 in ref. 17, p. 261). The tracer-labeled ROIs were transposed to the same space as the DWI data (Fig. 1), and the agreement between tracer results and tractography results was estimated in terms of sensitivity and specificity (30, 31). Additionally, to identify the tractography technique and/or the combination of user-defined tractography parameters that yielded the best overall combination of specificity and sensitivity, we computed the Youden index (J) [(specificity + sensitivity) −1] (32), a widely used measure in the cost-benefit analysis of diagnostic decision-making. The ideal tractography technique is expected to have maximum sensitivity (i.e., zero false negatives) without any decrease in specificity (i.e., zero false positives) and, therefore, a Youden index value of 1.

Effect of Diffusion Model. To test the effect of the diffusion model on the anatomical specificity and sensitivity of tractography, we generated tracts based on four diffusion models including DTI and three HARDI-type methods: Q-ball (23), constrained spherical deconvolution (CSD) (22), and ball and stick (B&S) (33). We used the default settings for tracking parameters (an angular threshold of 45° for deterministic techniques and 80° for probabilistic tractography) and for the tracking algorithm [e.g., fiber assignment by continuous tracking (11)]. The main finding is that, across the diverse set of tractography methods, an increase in sensitivity generally came at the cost of a decrease in specificity. As is evident from Fig. 2, this pattern can be observed across the two seed ROI locations and is consistent in both the deterministic and probabilistic tractography implementations that we tested. Among the deterministic techniques, the CSD approach showed the highest sensitivity across the different ROIs but also showed the lowest specificity. In contrast, the sensitivity and specificity of the Q-ball and DTI techniques varied widely between the two seed locations.

Although the tractography results of deterministic techniques are presented without any visualization thresholds (i.e., all identified tracts are displayed), the results of the B&S approach typically are visualized using an arbitrary threshold (i.e., only voxels that reach a given visitation frequency are displayed). To conduct a comprehensive assessment of all techniques, we compared the sensitivity and specificity of the B&S technique with three thresholds: (*i*) without applying a threshold (B&S-UT); (*ii*) by applying a threshold typically used in published studies of ex vivo brains (e.g., ref. 34) (B&S-ST); and (*iii*) by applying an optimal visualization threshold (B&S-OT) derived by generating the receiver operating characteristic (ROC) curve of all possible thresholds (*SI Materials and Methods*). As shown in Fig. 2, when no threshold is applied, the B&S approach offers maximum sensitivity but very poor specificity. Although applying a threshold for the B&S method reduces the false-positive rate dramatically, the use of a standard threshold (B&S-ST) results in the lowest sensitivity compared with all other methods, specifically for ROI-PCG, while offering the highest specificity. In contrast, the threshold derived from the ROC analysis (B&S-OT), which can be performed only if some ground truth data are available, provides a higher level of sensitivity with only a marginal decrease in specificity for ROI-PCG and ROI-V4v. In fact, a comparison of the Youden J values indicates that of the four tractography methods, the B&S-OT provides the optimal tradeoff between specificity and sensitivity across the two ROIs. Finally, although HARDI models generally are thought to have higher sensitivity than DTI (35), lower sensitivity is observed in the HARDI methods (Q-ball and B&S-ST in ROI-PCG and B&S-ST in ROI-V4v).

Effect of Angular Threshold. The angular threshold is the maximum bending angle allowed for a tract trajectory. This constraint prevents tracking algorithms from reconstructing fiber

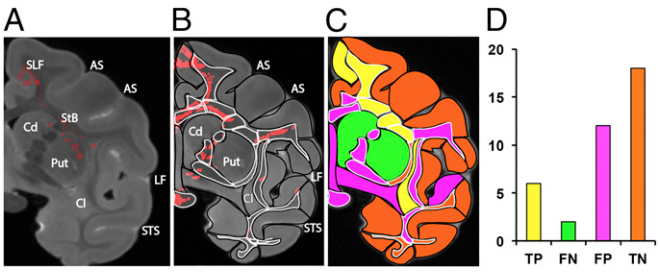

Fig. 1. A representative illustration of the procedure used for assessing the anatomical accuracy of various diffusion tractography approaches. (*A*) The red dots indicate the topography of axonal pathways following the injection of an anterograde tracer in the left PCG, redrawn from the reference atlas (data from ref. 17, p. 324) to the corresponding coronal slice of the DWI volume, which was rotated and resliced to match the histology slices used in the reference atlas. (*B*) For each DWI slice that was anatomically matched with the histology slice from the reference atlas, the gray matter (black lines) and white matter (white lines) regions were manually segmented and parcellated into discrete ROIs. The red dots in *B* represent the location of axonal pathways as determined with the Q-ball tractography method based on the left PCG as the seed region and an angular threshold of 80°. (*C*) Using the gray and white matter ROIs illustrated in *B* as a grid, the agreement between tracer and tractography results was computed for each slice. The colors indicate ROIs that were categorized as true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN). (*D*) Histogram showing the TP, FN, FP, and TN for the specific slice, which were used to compute the specificity [TN/(TN + FP)] and sensitivity [TP/(TP + FN)] values for the single slice shown here. AS, arcuate sulcus; Cd, caudate; Cl, claustrum; LF, lateral fissure; Put, putamen; SLF, superior longitudinal fasciculus; StB, striatal bundle; STS, superior temporal sulcus.

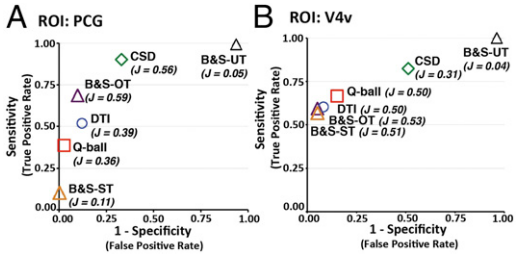

Fig. 2. The specificity and sensitivity of diffusion tractography techniques differ by diffusion model and location of the seed ROI. Specificity and sensitivity for seed ROI-PCG (*A*) and seed ROI-V4v (*B*). For all four types of diffusion models tested, the seed ROI was a sphere with a radius of 10 voxels, and the default angular threshold was used for deterministic (45°) and probabilistic (80°) tractography techniques. The Youden index value (J), which summarizes the performance of each tractography technique, is noted.

trajectories that are obviously anatomically incorrect. In general, the angular threshold is 45° for most deterministic tractography implementations and 80° for most probabilistic methods. Here, we explored the effect of four angular thresholds (30°, 45°, 60°, and 80°) on the specificity and sensitivity of the various tractography techniques. The key finding from our investigation is that, across all tractography methods, sensitivity increases but specificity decreases as angular threshold increases (Fig. 3). Our results suggest that for deterministic tractography techniques, a generic angular threshold such as 45° cannot be assumed because the optimal levels of specificity and sensitivity are dependent on the type of deterministic tractography method used, as well as on the location of the seed ROI. As shown in Fig. 3 *A–C* in the case of ROI-PCG, the optimal specificity and sensitivity are observed at 60° for DTI and Q-ball but at 45° for CSD. However, as is evident from Fig. 3 *E–G*, in the case of ROI-V4v, optimal specificity and sensitivity are observed at 45° for DTI and Q-ball and at 30° for CSD-based tractography.

It is noteworthy that the pattern of the effect of changes in angular threshold on specificity and sensitivity differs for the probabilistic tractography based on the B&S model. In particular, compared with all the deterministic tractography techniques, the probabilistic technique based on a data-driven estimation of the optimal visualization threshold (B&S-OT) shows the smallest decrease in specificity as the angular threshold increases from 30° to 80° (Fig. 3 *D* and *H*). However, unlike the deterministic tractography techniques, only a marginal increase in sensitivity can be observed for ROI-PCG, and very little change in sensitivity can be detected for ROI-V4v (Fig. 3*H*). Indeed, even at the extreme angular threshold of 80°, the Youden index does not increase beyond 0.59.

Composition of the Seed ROI. With the anterograde tracer technique, the radioactively labeled tracer is absorbed selectively by the cell bodies (i.e., gray matter) and transported to the axon terminals. In contrast, with DWI data from a standard clinical scanner, it generally is difficult for tractography algorithms to generate trajectories in gray matter because of its low anisotropy. Therefore, most diffusion tractography techniques require the seed ROIs to be dilated to include voxels containing white matter (e.g., ref. 36). However, an arbitrary increase in ROI size can result in structural connectivity maps that are not anatomically accurate. Techniques such as probabilistic tractography based on the B&S model have been reported to overcome this limitation, partly by using a liberal angular threshold of 80° (37). We therefore set the angular threshold to 80° for all four tractography techniques and tested their anatomical accuracy using seed ROIs that were restricted to gray matter. As shown in Fig. 4, when the ROI contained predominantly gray matter voxels, DTI showed the lowest sensitivity for ROI-PCG, whereas Q-ball showed the lowest sensitivity for ROI-V4v. Among the deterministic techniques, CSD showed the highest sensitivity across the two ROIs. For the B&S probabilistic technique, the highest sensitivity and the lowest specificity were observed when a visualization threshold was not imposed (B&S-UT). Applying the standard threshold (B&S-ST) resulted in the lowest sensitivity compared with all other methods for both ROIs but also offered the highest specificity. However, of the four tractography techniques, B&S-OT provided the optimal levels of sensitivity and specificity. That said, a comparison of the Youden J values for tractography results based on gray matter ROIs versus tractography results based on ROIs dilated to include white matter voxels shows dramatic changes in Youden J values for all three deterministic techniques (Table S1) as well as for the B&S probabilistic technique when the results are unthresholded or are thresholded using standard recommendations. In summary, sensitivity generally is reduced when the seed ROI is predominantly within gray matter, and although dilating the ROI improves sensitivity with deterministic tractography techniques, it does so at the cost of dramatically reduced specificity. In addition, the probabilistic technique appears to be less susceptible to changes in the composition of an ROI, but only if an optimized threshold can be derived and used.

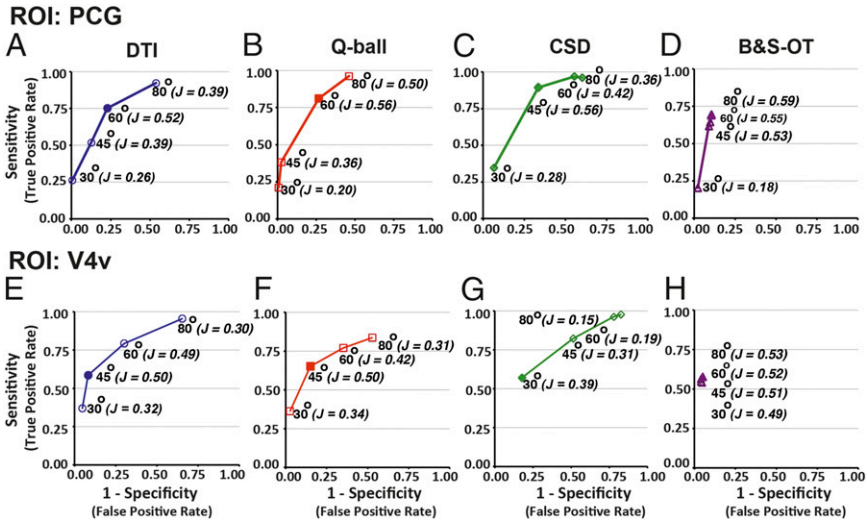

Fig. 3. The effect of changing the angular threshold on the specificity and sensitivity of results from the four tractography techniques across the two seed ROIs. (*A–H*) For both ROIs, the Youden index value (J) for each angular threshold is noted, and the threshold that yields the most optimal J value is depicted by a filled symbol. Note that for the B&S method, the specificity and sensitivity are presented only for the optimal threshold. See Figs. S1 and S2 for the ROC curves for each angular threshold.

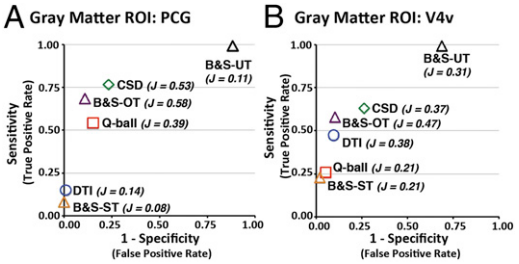

Fig. 4. Sensitivity and specificity of the four tractography techniques in seed location PCG (*A*) and V4v (*B*) when the seed ROI was restricted within gray matter. For all four types of diffusion models tested, the angular threshold was set to 80°. See Fig. S3 for a summary of the true positives and false negatives for the two seed regions.

Discussion

Our primary objective was to test the hypothesis that DWI data of exceptional quality and sophisticated DWI tractography techniques are sufficient to generate an anatomically accurate map of axonal connections. Using ultra-high-resolution ex vivo DWI in the macaque and previously published tracer results, we undertook a comprehensive assessment of the sensitivity and specificity of a variety of widely used tractography techniques. The main finding of our study is that a tractography technique that shows high sensitivity (a high rate of true positives) most likely will show low specificity (a high rate of false positives). In addition, the anatomical accuracy of tractography techniques was found to be highly dependent on a number of technical parameters, such as the type of diffusion model, the angular threshold, and the composition of the seed ROI. Moreover, our results indicate that the choice of parameters that produces the best combination of sensitivity and specificity varies for different pathways. Overall, even in ideal experimental conditions, the anatomical accuracy of DWI tractography is suboptimal.

This conclusion initially may seem at odds with the few published studies that have reported validation of tractography with histological data in the macaque (34, 38). However, in these studies the validation was based primarily on qualitative observations. Studies that assessed the accuracy of DWI tractography by quantifying the overlap fraction between tracer-labeled regions and tractography (39, 40) found that DWI tractography missed some of the tracer-labeled target regions (i.e., false negatives). However, these previous studies differ from the present study in three ways. First, they were focused primarily on assessing sensitivity rather than specificity (34, 38). Our study shows that a very high level of sensitivity can be achieved by changing certain tracking parameters such as the angular threshold. However, the gain in sensitivity comes at the cost of reduced specificity. Second, in the previous studies, the tractography results were constrained either by waypoints (34, 39) or orientation-specific ROIs (38). Using a priori anatomical knowledge to constrain tractography is likely to reduce the occurrence of false-positive trajectories but precludes an objective assessment of the ability of the technique to reveal the trajectory of fiber pathways that are currently unknown. Finally, previous studies were less general and systematic in assessing the effects of a broad range of tractography algorithms and implementations, because they focused on validating one specific implementation of tractography using specific parameters. In the present study we tested the accuracy of different types of tractography implementations ranging from the simple diffusion tensor model to sophisticated HARDI diffusion models that are considered capable of resolving complex fiber crossings. However, even the most sophisticated tractography methods tested here did not consistently show superior sensitivity and specificity.

Implications. The major implication of our study is that a diffusion tractography technique that produces anatomically accurate results remains an elusive goal even with DWI data of exceptional quality. The ex vivo DWI data used here are of such high quality that, by our estimation, acquiring data of the same SNR and resolution from a human brain in a standard clinical scanner, in vivo, would require thousands of hours of scan time. We believe our results highlight an inherent limitation of DWI tractography: inferring fiber direction information from a water diffusion displacement profile is fundamentally a complex, underdetermined inverse problem that cannot be solved. Our results also suggest that this problem is unlikely to be overcome by using sophisticated HARDI tracking algorithms or by acquiring high-resolution (angular and spatial) DWI data. In this respect, our findings are consistent with an ex vivo DWI tractography study of the human optic chiasm that showed that DWI tractography fails to identify the ipsilateral and contralateral axonal pathways that branch at the optic chiasm, even at a resolution of 156 μm (in-plane) (41). Despite this intrinsic limitation, tractography is likely to remain a widely used tool in neuroscience, because currently it is the only noninvasive method that allows visualization of white matter pathways in vivo. So how can our findings help improve the use of tractography?

Our results indicate that none of the four tractography techniques we assessed can be considered superior for all applications of tractography. Future improvements in the accuracy of diffusion tractography will require innovations in MRI hardware, sequence design, data acquisition strategies, and tractography algorithms (28, 42). Although such advances will lead to incremental improvements in the overall accuracy, they may not overcome the inherent ambiguities in inferring long-range axonal connectivity based on local diffusion displacement profiles. One suggestion, therefore, is to select the tractography method, or combination of methods, most appropriate for a specific objective. For example, if the objective is to reduce the possibility of identifying spurious pathways, a tractography method with better specificity, such as DTI, Q-ball, or B&S probabilistic tractography (using a conservative threshold), should be used. Alternatively, if the objective is to reduce the likelihood of missing salient pathways, a tractography technique with relatively high sensitivity, such as CSD and B&S probabilistic tractography (using a liberal visualization threshold), would be more appropriate. For example, to avoid inadvertent transection of critical fibers of passage, as in the case of palliative surgery for brain tumors, a tractography technique with low specificity but high sensitivity would be appropriate.

It is important to note that our results do not invalidate the utility of DWI tractography for visualizing major fasciculi. Indeed, many of the major fasciculi segmented using tractography are anatomically consistent with dissection studies in the human (43, 44). Instead, our results underscore the fact that DWI tractography alone will not be sufficient to build an anatomically accurate map of the human brain connectome. Moreover, given the inherent ambiguity of DWI tractography results, it is crucial to support inferences about the organization of pathways in the human with data from animal models, using converging methods. Granted, the use of animal models is somewhat limited because of interspecies differences in neuroanatomy (45, 46). However, given the similarities in the connectivity profile between the monkey and human brain (47, 48), macaque models will continue to serve as an indispensable tool for explicating general organizational principles of brain networks such as characterization of the directional transitions that axons make when entering or exiting gray matter (35, 49, 50). Insights from such studies may help improve the accuracy of tractography techniques.

Caveats. Our study has some potential limitations. One concern is that, because of the use of anterograde tracers, the tracer data used here document only the afferent projections from the site of the tracer injection. Connections visualized with DWI tractography, however, do not differentiate between afferent and efferent projections. Thus, some of the false positives identified in the present analysis could possibly be true positives; as a result, the specificity of all tractography techniques would be underestimated. However, studies that used retrograde tracers in the macaque (51) and other primates (52, 53) have found that the cortico-cortical and cortico-thalamic regions identified as being directly connected to the injection sites PCG and V4v tend to be bidirectional. Moreover, our own results suggest that this aspect does not contribute to a large bias in the rate of false positives. As is evident from Fig. 2*A*, the rate of false positives for Q-ball reaches values very close to 0, and sensitivity is greater than chance, which would not be possible in the presence of false positives caused by retrograde connections. This pattern can be observed in the performance of B&S-OT and DTI for ROI-V4v as well (Fig. 2*B*).

Another potential issue is that the tracer and DWI data are not from the same animals, and therefore factors such as interindividual variability in neuroanatomy and mismatches in the placement of the ROIs used to assess agreement between DWI and tracer data could negatively impact the measured accuracy of the tractography methods tested here. With regard to variability in the connectional neuroanatomy, tracer studies involving multiple monkeys suggest there is very little interindividual variability in the topography of connections (48, 54). Morphological variability can be present, but we have mitigated the potential effects of this confound by using careful registration strategies. The DWI data were 3D-registered to match the tracer atlas at each slice location optimally, ensuring local correspondence between anatomical landmarks such as major sulci and the location of subcortical nuclei and commissural pathways. To minimize the effects of potential mismatches in the extent and location of the ROIs used to assess agreement between DWI and tracer data, we defined a grid of relatively large, anatomically meaningful ROIs without requiring agreement at a fine level (e.g., on a voxel-by-voxel or a fixed millimeter-scale grid). Thus, a tracer ROI that was identified at the head or tail of the caudate nucleus would be transferred to the same anatomical region in the MRI slice even if the whole brain morphometry was slightly different in the two specimens. Our own results suggest that differences in neuroanatomy do not bias the performance of the tractography methods tested here. In fact, any systematic bias in the registration of the two datasets or anatomical parcellation should affect the measurement of accuracy of all the tractography techniques tested here, for all possible settings of the algorithm, which was not the case (Fig. S3).

Another aspect that potentially could impact tractography results is the choice of b-values used in the DWI acquisition. In the present study, we used a single b-value of 4,800 s/mm^2 that previously was found to be within the optimal range (~2,000 to ~8,000 s/mm^2) for most tractography methods in fixed brain specimens (55). However, Dyrby et al. (55) found that some HARDI methods, such as Q-ball showed optimal tractography results with higher b-values. Given that the optimal b-value is method dependent, the performance of specific tractography methods can be expected to vary as a function of the b-value. However, we expect the sensitivity-specific tradeoff observed across the various tractography methods will persist regardless of the b-value used for data acquisition.

Finally, we note that each of the tractography methods assessed here contains many more user-tunable parameters than those examined in the present study. For example, one could use the probabilistic version of the CSD method instead of the deterministic version to improve the robustness of the tractography results (56). Furthermore, adjusting parameters within each method, such as the harmonic degree used for CSD, the step size used in the tractography algorithm, and the type of interpolation performed, among others, would result primarily in a shift of the sensitivity/specificity values along the ROC curves, although some effects on overall performance may be possible. A systematic examination of the effect of all possible parameters is beyond the scope of this work; however, the methodology we adopted here offers an objective approach for assessing the effect of additional parameters in the future and potentially fine-tuning their optimal value.

Capitalizing on the Strengths of DWI. Although the results of our study highlight a fundamental limitation in using DWI tractography to map brain connectivity accurately, DWI will continue to be an extremely valuable tool in neuroscience. The core strength of DWI is that it uses measurements of the displacement of water molecules to reveal the local microstructural features of biological tissue that typically are hidden in standard structural scans (Fig. S4). Specifically, DWI makes it possible to visualize the anisotropy of water diffusion in brain tissue and to derive quantitative measures related to the microstructural and architectural features of white matter and other types of brain tissue (57). This feature makes DWI an indispensable research tool for investigating brain pathology. In addition, DWI measurements of local tissue microstructure are being used to explore the relationship between individual differences in white matter architecture and behavior and to measure changes in the brain as a result of developmental maturation, learning, and aging. Moving forward, it is worthwhile to recognize that the diffusion displacement profile, mathematically described as the diffusion propagator (58, 59), contains not only orientational information but also information pertaining to water compartments and the exchange between those compartments. Therefore, more comprehensive characterization of the diffusion propagator can provide a series of novel quantitative measures (60) related to the microstructural properties of brain tissue. Such measures can further improve our ability to quantify changes in brain microstructure in healthy and clinical populations.

Materials and Methods

MRI Data Acquisition. The brain of one adult male Rhesus macaque monkey (*Macaca mulatta*) (4.5 y old, 5.95 kg) was used in this study (see *SI Materials and Methods* for details regarding sample preparation.) A standard linear birdcage volume coil (72 mm i.d.) was used to image the brain in a 30-cm magnet bore 7-Tesla scanner (Bruker BioSpin). The scanner was equipped with a custom-designed gradient coil (Resonance Research Inc.) capable of a maximum gradient strength of 440 mT/m on each axis and 120-μs ramp time. The diffusion data were acquired with a 3D spin-echo diffusion-weighted EPI sequence (DtiEpi) available in Bruker Paravision software. The diffusion-encoding part comprised a pair of Stejskal–Tanner gradient pulses, with $\delta = 6$ ms and $\Delta = 14$ ms, and the b-value was set at 4,800 s/mm^2, which has been demonstrated to be sufficient to model multiple fiber populations in ex vivo specimens (55). A diffusion-weighting gradient table of 121 directions on vertices of a tessellated icosahedral hemisphere was used, and seven additional image volumes were collected with $b = 0$. A 3D EPI scan was used to obtain isotropic high-resolution images with high SNR. To avoid EPI ghosting artifacts, even and odd echoes were separately reconstructed into magnitude images before the two sets of images were averaged together (61). Echo time was reduced by acquiring the EPI train in 16 segments. Partial Fourier reconstruction was used in the EPI phase-encoding direction to reduce the echo time further, with an oversampling factor of 1.3. The 3D image matrix size was 278 × 256 × 238, resulting in an isotropic spatial resolution of 250 μm. Scan repetition time was 500 ms; echo time was 34 ms. Each diffusion direction took about 32 min to acquire, and the entire DWI acquisition took ~71 h. All procedures followed the *Guide for the Care and Use of Laboratory Animals* (62) and were approved by the National Institute of Mental Health Animal Care and Use Committee.

DWI Preprocessing and Tractography. The data were preprocessed using the TORTOISE software package (63) and were corrected for eddy current distortions and motion-like artifacts caused by frequency drifts. The preprocessed DWI data were processed further through specific software pipelines to compute the voxelwise diffusion displacement profile, which forms the basis for all tractography methods. For all the tractography methods (DTI, Q-ball, CSD, and B&S), the only stopping criterion was the angular threshold, which was set to 30°, 45°, 60°, 80°, or its equivalent in terms of radius of curvature. Based on the tracer datasets, the location and the size of the seed ROIs were defined, and the

tractography results were visualized. Finally, the agreement between the tracer results and tractography results was computed using in-house software (for details see *SI Materials and Methods*).

ACKNOWLEDGMENTS. We thank Drs. Peter Basser, Elizabeth Hutchinson, Alexandru Avram, Danping Liu, and Neda Sadeghi for useful comments and discussion. Special thanks go to Dr. Afonso Silva for providing the MRI scanner time and instrument support. This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development and National Institute of Mental Health Intramural Programs. Salary support for C.T. and M.O.I. was provided by funding from the Department of Defense in the Center for Neuroscience and Regenerative Medicine.

1. Schmahmann JD, Pandya DN (2007) Cerebral white matter—historical evolution of facts and notions concerning the organization of the fiber pathways of the brain. *J Hist Neurosci* 16(3):237–267.
2. Le Bihan D, et al. (1986) MR imaging of intravoxel incoherent motions: Application to diffusion and perfusion in neurologic disorders. *Radiology* 161(2):401–407.
3. Merboldt K-D, Hanicke W, Frahm J (1985) Self-diffusion NMR imaging using stimulated echoes. *J Magn Reson (1969)* 64(3):479–486.
4. Taylor DG, Bushell MC (1985) The spatial mapping of translational diffusion coefficients by the NMR imaging technique. *Phys Med Biol* 30(4):345–349.
5. Basser PJ, Mattiello J, Le Bihan D (1994) Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 103(3):247–254.
6. Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G (1996) Diffusion tensor MR imaging of the human brain. *Radiology* 201(3):637–648.
7. Sundgren PC, et al. (2004) Diffusion tensor imaging of the brain: Review of clinical applications. *Neuroradiology* 46(5):339–350.
8. Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A (2000) In vivo fiber tractography using DT-MRI data. *Magn Reson Med* 44(4):625–632.
9. Conturo TE, et al. (1999) Tracking neuronal fiber pathways in the living human brain. *Proc Natl Acad Sci USA* 96(18):10422–10427.
10. Jones DK, Simmons A, Williams SC, Horsfield MA (1999) Non-invasive assessment of axonal fiber connectivity in the human brain via diffusion tensor MRI. *Magn Reson Med* 42(1):37–41.
11. Mori S, Crain BJ, Chacko VP, van Zijl PC (1999) Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* 45(2):265–269.
12. Poupon C, et al. (2000) Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *Neuroimage* 12(2):184–195.
13. Jbabdi S, Johansen-Berg H (2011) Tractography: Where do we go from here? *Brain Connect* 1(3):169–183.
14. Jones DK (2008) Studying connections in the living human brain with diffusion MRI. *Cortex* 44(8):936–952.
15. Catani M, de Schotten MT (2012) *Atlas of Human Brain Connections* (Oxford Univ Press, New York).
16. Johansen-Berg H, Behrens TE (2006) Just pretty pictures? What diffusion tractography can add in clinical neuroscience. *Curr Opin Neurol* 19(4):379–385.
17. Schmahmann JD, Pandya D (2009) *Fiber Pathways of the Brain* (Oxford Univ Press, New York).
18. Pierpaoli C, et al. (2001) Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *Neuroimage* 13(6 Pt 1):1174–1185.
19. Farquharson S, et al. (2013) White matter fiber tractography: Why we need to move beyond DTI. *J Neurosurg* 118(6):1367–1377.
20. Cook P, et al. (2006) Camino: Open-source diffusion-MRI reconstruction and processing. *14th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, p 2759.
21. Dell'Acqua F, et al. (2007) A model-based deconvolution approach to solve fiber crossing in diffusion-weighted MR imaging. *Biomedical Engineering. IEEE Transactions on* 54(3):462–472.
22. Tournier JD, Calamante F, Connelly A (2007) Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35(4):1459–1472.
23. Tuch DS, et al. (2002) High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med* 48(4):577–582.
24. Wedeen VJ, Hagmann P, Tseng W-YI, Reese TG, Weisskoff RM (2005) Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn Reson Med* 54(6):1377–1386.
25. Pierpaoli C (2011) Artifacts in Diffusion MRI. *Diffusion MRI: Theory, Methods and Application,* ed Jones D (Oxford Publications, New York), pp 303–318.
26. İrfanoğlu MO, Walker L, Sarlls J, Marenco S, Pierpaoli C (2012) Effects of image distortions originating from susceptibility variations and concomitant fields on diffusion MRI tractography results. *Neuroimage* 61(1):275–288.
27. Uğurbil K, et al.; WU-Minn HCP Consortium (2013) Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *Neuroimage* 80:80–104.
28. Setsompop K, et al. (2013) Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *Neuroimage* 80:220–233.
29. Van Essen DC, et al.; WU-Minn HCP Consortium (2012) The Human Connectome Project: A data acquisition perspective. *Neuroimage* 62(4):2222–2231.
30. Seehaus AK, et al. (2012) Histological validation of DW-MRI tractography in human postmortem tissue. *Cereb Cortex* 23(2):442–450.
31. Iturria-Medina Y, et al. (2011) Brain hemispheric structural efficiency and interconnectivity rightward asymmetry in human and nonhuman primates. *Cereb Cortex* 21(1):56–67.
32. Youden WJ (1950) Index for rating diagnostic tests. *Cancer* 3(1):32–35.
33. Behrens TE, Berg HJ, Jbabdi S, Rushworth MF, Woolrich MW (2007) Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *Neuroimage* 34(1):144–155.
34. Jbabdi S, Lehman JF, Haber SN, Behrens TE (2013) Human and monkey ventral prefrontal fibers use the same organizational principles to reach their targets: Tracing versus tractography. *J Neurosci* 33(7):3190–3201.
35. Sotiropoulos SN, et al.; WU-Minn HCP Consortium (2013) Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *Neuroimage* 80:125–143.
36. Gschwind M, Pourtois G, Schwartz S, Van De Ville D, Vuilleumier P (2012) White-matter connectivity between face-responsive regions in the human brain. *Cereb Cortex* 22(7):1564–1576.
37. Behrens TE, et al. (2003) Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nat Neurosci* 6(7):750–757.
38. Schmahmann JD, et al. (2007) Association fibre pathways of the brain: Parallel observations from diffusion spectrum imaging and autoradiography. *Brain* 130(Pt 3):630–653.
39. Dyrby TB, et al. (2007) Validation of in vitro probabilistic tractography. *Neuroimage* 37(4):1267–1277.
40. Dauguet J, et al. (2007) Comparison of fiber tracts derived from in-vivo DTI tractography with 3D histological neural tract tracer reconstruction on a macaque brain. *Neuroimage* 37(2):530–538.
41. Roebroeck A, et al. (2008) High-resolution diffusion tensor imaging and tractography of the human optic chiasm at 9.4 T. *Neuroimage* 39(1):157–168.
42. Mangin J-F, et al. (2013) Toward global tractography. *Neuroimage* 80:290–296.
43. Bürgel U, et al. (2006) White matter fiber tracts of the human brain: Three-dimensional mapping at microscopic resolution, topography and intersubject variability. *Neuroimage* 29(4):1092–1105.
44. Lawes IN, et al. (2008) Atlas-based segmentation of white matter tracts of the human brain using diffusion tensor tractography and comparison with classical dissection. *Neuroimage* 39(1):62–79.
45. Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. *Proc Natl Acad Sci USA* 105(49):19514–19519.
46. Saleem KS, Price JL, Hashikawa T (2007) Cytoarchitectonic and chemoarchitectonic subdivisions of the perirhinal and parahippocampal cortices in macaque monkeys. *J Comp Neurol* 500(6):973–1006.
47. Margulies DS, et al. (2009) Precuneus shares intrinsic functional architecture in humans and monkeys. *Proc Natl Acad Sci USA* 106(47):20069–20074.
48. Saleem KS, Miller B, Price JL (2014) Subdivisions and connectional networks of the lateral prefrontal cortex in the macaque monkey. *J Comp Neurol* 522(7):1641–1690.
49. Leuze CW, et al. (2014) Layer-specific intracortical connectivity revealed with diffusion MRI. *Cereb Cortex* 24(2):328–339.
50. McNab JA, et al. (2013) Surface based analysis of diffusion orientation for identifying architectonic domains in the in vivo human cortex. *Neuroimage* 69:87–100.
51. Ungerleider LG, Galkin TW, Desimone R, Gattass R (2008) Cortical connections of area V4 in the macaque. *Cereb Cortex* 18(3):477–499.
52. Stepniewska I, Preuss TM, Kaas JH (1994) Thalamic connections of the primary motor cortex (M1) of owl monkeys. *J Comp Neurol* 349(4):558–582.
53. Burman KJ, Bakola S, Richardson KE, Reser DH, Rosa MG (2014) Patterns of cortical input to the primary motor area in the marmoset monkey. *J Comp Neurol* 522(4):811–843.
54. Webster MJ, Bachevalier J, Ungerleider LG (1994) Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb Cortex* 4(5):470–483.
55. Dyrby TB, et al. (2011) An ex vivo imaging pipeline for producing high-quality and high-resolution diffusion-weighted imaging datasets. *Hum Brain Mapp* 32(4):544–563.
56. Tournier J, Calamante F, Connelly A (2012) MRtrix: Diffusion tractography in crossing fiber regions. *Int J Imaging Syst Technol* 22(1):53–66.
57. Budde MD, Frank JA (2010) Neurite beading is sufficient to decrease the apparent diffusion coefficient after ischemic stroke. *Proc Natl Acad Sci USA* 107(32):14472–14477.
58. Callaghan PT (1993) *Principles of Nuclear Magnetic Resonance Microscopy* (Oxford Univ Press, New York).
59. Stejskal E (1965) Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *J Chem Phys* 43:3597.
60. Özarslan E, et al. (2013) Mean apparent propagator (MAP) MRI: A novel diffusion imaging method for mapping tissue microstructure. *Neuroimage* 78:16–32.
61. Yang QX, Posse S, Le Bihan D, Smith MB (1996) Double-sampled echo-planar imaging at 3 tesla. *J Magn Reson B* 113(2):145–150.
62. Committee on Care and Use of Laboratory Animals (1996) *Guide for the Care and Use of Laboratory Animals* (National Academies Press, Washington, DC).
63. Pierpaoli C, et al. (2010) TORTOISE: An integrated software package for processing of diffusion MRI data. *18th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, p. 1597.