Magnetic Resonance Imaging 55 (2019) 7–25

Contents lists available at ScienceDirect

Magnetic Resonance Imaging

journal homepage: www.elsevier.com/locate/mri

Original contribution

Anatomical accuracy of standard-practice tractography algorithms in the motor system - A histological validation in the squirrel monkey brain

Kurt G. Schilling[a,b,*], Yurui Gao[a,b], Iwona Stepniewska[c], Vaibhav Janve[a,b], Bennett A. Landman[a,b,d,e], Adam W. Anderson[a,b,d]

[a] Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, USA
[b] Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA
[c] Department of Psychology, Vanderbilt University, Nashville, TN, USA
[d] Radiology and Radiological Sciences, Vanderbilt University, Nashville, TN, USA
[e] Department of Electrical Engineering, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Keywords:
Validation
Diffusion magnetic resonance imaging
Histology
Tractography
Dispersion
HARDI
Accuracy

ABSTRACT

For two decades diffusion fiber tractography has been used to probe both the spatial extent of white matter pathways and the region to region connectivity of the brain. In both cases, anatomical accuracy of tractography is critical for sound scientific conclusions. Here we assess and validate the algorithms and tractography implementations that have been most widely used - often because of ease of use, algorithm simplicity, or availability offered in open source software. Comparing forty tractography results to a ground truth defined by histological tracers in the primary motor cortex on the same squirrel monkey brains, we assess tract fidelity on the scale of voxels as well as over larger spatial domains or regional connectivity. No algorithms are successful in all metrics, and, in fact, some implementations fail to reconstruct large portions of pathways or identify major points of connectivity. The accuracy is most dependent on reconstruction method and tracking algorithm, as well as the seed region and how this region is utilized. We also note a tremendous variability in the results, even though the same MR images act as inputs to all algorithms. In addition, anatomical accuracy is significantly decreased at increased distances from the seed. An analysis of the spatial errors in tractography reveals that many techniques have trouble properly leaving the gray matter, and many only reveal connectivity to adjacent regions of interest. These results show that the most commonly implemented algorithms have several shortcomings and limitations, and choices in implementations lead to very different results. This study should provide guidance for algorithm choices based on study requirements for sensitivity, specificity, or the need to identify particular connections, and should serve as a heuristic for future developments in tractography.

1. Introduction

Diffusion MRI fiber tractography is widely used to probe the structural connectivity of the brain, with a range of applications in both clinical and basic neuroscience [1,2]. However, these techniques are subject to a number of serious pitfalls and limitations which may limit the anatomical accuracy of the reconstructed pathways [3,4]. In addition, the large number of diffusion reconstruction algorithms and tracking strategies that exist are likely to result in different "tracts", with varying levels of accuracy. As utilization of fiber tractography continually increases, it is necessary to validate these techniques in order to gain insight into the conditions under which they succeed, and more importantly, where they fail.

One approach to validation is through classical tracer injection techniques in animal models, followed by histological analysis to defined the "ground truth" pathways for subsequent comparisons with diffusion tractography. Traditionally, validating the faithfulness of tractography against tracers takes one of two forms. First, some metric of spatial overlap of the tract versus the tracer can be computed, which evaluates the overall layout or spatial extent of the tract. Second, many studies evaluate connectivity measures, disregarding *how* tracts reach their destinations, with a focus on the strength of the connections between different regions of the brain.

Connection strengths estimated from tractography have been compared with invasive tracer data accumulated in existing atlases or databases, for example the Markov-Kennedy [5] or CoCoMac [6]

* Corresponding author at: Vanderbilt University Institute of Imaging Science, Vanderbilt University, Nashville, TN, USA.
E-mail address: kurt.g.schilling.1@vumc.org (K.G. Schilling).

https://doi.org/10.1016/j.mri.2018.09.004
Received 11 June 2018; Received in revised form 6 August 2018; Accepted 6 September 2018

databases for the macaque, or the Allen Brain Atlas for the mouse [7]. These studies have provided encouraging results, finding moderate to high positive correlations between tractography and connection strengths [8–10], suggesting that the number of reconstructed streamlines is correlated with the strength of connections between brain regions. However, tractography becomes less reliable for longer pathways [10], and results are heavily dependent on decisions made during the tracking process (i.e. the seeding strategy). The use of large-scale tracer databases has the advantage of assessing connectivity of a large number of pathways across many cortical areas, however, they have several disadvantages. Most notably, tracer injection and MRI are typically not employed on the same animal (with few exceptions [11]). Not only can pathway connection strength vary between animals, but variance in brain geometry between injected and scanned animals could lead to mismatches in identifying the location of the injection regions in the subject of interest, together compromising the fidelity of the "ground truth" to which tractography is being compared.

Alternatively, a number of studies have investigated the voxel-wise spatial overlap of histologically-defined white matter trajectories with those from tractography. Validating these measures gives confidence in the ability of tractography to segment specific white matter pathways (with subsequent analysis typically taking some quantitative measure along these pathways). For example, Schmahmann et al. [12] compare one implementation of tractography (diffusion spectrum imaging) to histological tracing and conclude that tractography is able to replicate the major features and geometrical organization of a number of association pathways. Improving upon this, in a series of studies on the macaque brain, Dauget et al. [13,14] register histological sections of labeled fiber tracts in 3D to diffusion tensor imaging (DTI) tractography data. They find a range in spatial agreement, with a range of Dice overlap coefficients (0.2–0.75) dependent on the pathway of interest and various tractography parameters, and note that DTI has difficulties when tracts cross or divide, an issue now referred to as the "crossing fiber" problem.

Building upon these studies, the goal of the present work is to systematically characterize the anatomical accuracy of diffusion fiber tractography – both the spatial extent and tract connections – and to do this on both the scale of individual voxels as well on a larger domain over anatomical regions of interest. To achieve this goal, we utilize the squirrel monkey brain, and compare tractography results directly to registered high-resolution tracer data from the same animal. We aim here to evaluate the algorithms most commonly employed in the literature (all of which are implemented in open-source software packages) in order to reveal the successes and shortcomings of the majority of studies utilizing diffusion tractography to date. In addition to measures of overlap and connectivity for each algorithm, we further assess the effects of user-defined algorithm choices (reconstruction algorithm, seeding strategy, tracking logic), distance from seed point, and effects of probabilistic thresholding on the fidelity of resulting tractograms. The focus of this work is on tractography of the pathways in the motor system. This is because the organization and anatomical connections of this system are well understood [15], and the motor system is a frequent target of tractography as it is particularly relevant for a variety of disabilities or pathologies including stroke [16,17], multiple sclerosis [18,19], Parkinson's disease [20,21], cerebral palsy [22,23], and tumor removal surgeries [24,25], among others. Herein, we investigate the spatial errors in these tractography algorithms, asking where in the brain these algorithms typically fail, and assess potential reasons for this failure.

2. Methods

All animal procedures were approved by the Vanderbilt University Animal Care and Use Committee. Fig. 1 shows the methodology pipeline used in this study. Briefly, biotinylated dextran amine (BDA), a histological tracer, was injected into the primary motor cortex (M1) of two squirrel monkey brains. Afterwards, diffusion MRI was acquired on the ex vivo brains and diffusion fiber tractography performed using 40 different algorithms and/or tracking settings. These 40 tractograms resulted in both streamline locations (with the exception of two algorithms) and track density maps, which represent the number of streamlines traversing each voxel. During the brain sectioning digital photographs of the frozen block of the brains were taken to aid in registration of the modalities. Histological sections were processed to visualize BDA and imaged at high resolution. BDA was then segmented from these images in order to create BDA density maps which can be compared directly on a voxel-by-voxel basis to the tractograms and streamline density maps.

2.1. Tracer injection

For our study we chose BDA, a commonly used neuroanatomical tracer for studying neuronal pathways. BDA is transported in both anterograde and retrograde directions, yielding sensitive and detailed labeling of both axons and terminals, as well as neuronal cell bodies [26]. This tracer relies on axonal transport systems; thus BDA injection was performed prior to ex vivo scanning. Under general anesthesia, BDA (Molecular Probes Inc., Eugene, OR) was injected (as a 10% solution in phosphate buffer) into M1 cortex of the left-hemisphere following the procedures followed in previous studies [11,15]. Pressure injections of BDA were carried out using a 2 μl Hamilton syringe. Eight injections (1 μl/site) were made in order to cover a large M1 region representing the distal forelimb as identified by intracortical microstimulation. After each injection, the needle was left in the brain for 5–10 min and then retracted stepwise to avoid leakage of the tracer along the needle track. After surgery, the monkeys were allowed to recover, giving the tracer sufficient time to be transported along axons to all regions connected to the injected M1 cortex.

2.2. Diffusion MRI acquisition

After animal sacrifice, the brains were perfusion fixed with 4% paraformaldehyde preceded by rinse with physiological saline. The brain was removed from the skull and stored in buffered saline overnight. The next day, the brain was scanned on a 9.4 T Varian scanner using a quadrature birdcage volume coil (inner diameter = 63 mm), and immersed in PBS during scanning. Diffusion weighted imaging was performed using a pulsed gradient spin echo multi-shot spinwarp imaging sequence with full brain coverage at 300 μm isotropic resolution (TR = 4.6 s, TE = 42 ms, 32 gradient directions, b ≈ 1000 s/mm^2, image matrix = 192 × 128 × 115, 1 b0 image). With a brain volume of approximately 20 cm^3, this is roughly equivalent to high resolution protocol of a human brain scanned at ~1.2 mm isotropic. Bore temperature was monitored and maintained at 19–20 °C by circulating air through the bore. Acquisition for a single diffusion-weighted volume took approximately 10 min. The scan time was extended to 50 h in order to facilitate 9 and 10 signal averages, respectively. The b value used in this experiment was lower than is optimal for diffusion studies in fixed tissue [27,28] (approximate mean diffusivity of 0.45 × 10–3 mm^2/s [29], about half of that expected in vivo), due to hardware limitations. A low b value decreases the diffusion-related contrast-to-noise ratio (CNR) in the image data (upon which tractography ultimately relies), which has the same effect as higher image noise. To compensate for this shortcoming, we extended the scan time to 50 h, which yielded a CNR comparable to in vivo human studies (equivalent to an in vivo study with mean diffusivity = 0.8 × 10–3 mm^2/s and SNR > 50). Thus, although the b-value is lower than optimal, the angular contrast is increased, resulting in voxel-wise reconstructions consistent with expected anatomy. See Supplementary Fig. 1 for example orientation distributions derived from example CSD, QBI, and B&S methods in both single fiber, crossing fiber, and gray matter regions. Because the scan is ex vivo, and only