**Title**
Diffusion MR Image Processing Tools for Reliable Fiber Tracking Analyses: Neurosurgery and Radiation Oncology Applications



**Author**
Jordan, Kesshi Marin

**Publication Date**
2017

Peer reviewed|Thesis/dissertation

# Diffusion MR Image Processing Tools for Reliable Fiber Tracking Analyses: Neurosurgery and Radiation Oncology Applications

by

Kesshi Marin Jordan

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Bioengineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

[redacted]

This work is dedicated to the Neurosurgery Patients at UCSF and their Families...

To the many of you who have shared your journeys with me over the past several years: sifting through your data in search of clues has left me in awe of the resilience of the human brain and spirit. You have made me a better scientist and a better human being. Thank you.

# Acknowledgments

I would like to foremost thank my advisor, Roland, for creating such an engaging environment in which to do science. His genuine excitement in pursuit of discovery has brought together a fantastic group of researchers. Thank you for the guidance, opportunities, and friendship.

I would also like to thank Bob, Pratik, and Mitch for serving as additional mentors on my committee, as well as my professors at both UCSF and UC Berkeley: I am lucky to have such a variety of expertise to learn from. I would also like to acknowledge my rotation advisors, administrators, funding from the NDSEG Fellowship and NIH, and our wonderful research subjects at UCSF, without whom none of this work would have been possible.

I would also like to express deep appreciation for the most delightful group of labmates and collaborators anyone could hope for, especially: Anisha and Bago, who turned me into a programmer; Nico and Eduardo, who are so generous with their expertise and a force to reckon with on the soccer field; Regina, Esha, and Antje, from whom I have learned so much both within and beyond the lab; Antonella, who taught me tumor neuroanatomy and how to select a shoe, Marco, Joe, Seunggu, Monica, Alyssa, Gina, Alan, Malu, Jason, Elke, Stephanie, Olivier, Mike, Julia, Lucia, Vikram, Deanna, Heather, Eddie, Simone, Francesco, Tristan, James, Jennifer, Anette, Cathy, Nandita, Vanitha, and Ziba. I looked forward to coming to work every day for the last five years because of you.

I would also like to acknowledge my family: my mother, who shared with me her deep love of science and put her considerable intellect into raising four incredible children; my father, whose "practical genius" was the perfect balance to my mother's "super brain"; Alanna, whose fierce confidence I have endeavored to channel; Delaney, whose blend of brilliance & creativity and Rider, whose quiet intellect & honor I strive to emulate; and Gwen, my happiest and furriest sibling. I would like to especially thank my husband, Andrew, and dog, Brie for their unflagging support and unconditional love. Without you I would have missed out on countless adventures during the last five years. PhD students are not the most lovable creatures, so those that put up with them deserve a great deal of credit.

Last but not least, my deepest admiration and thanks to JMR, ADM, AEW, MCD, RBG, HRC et al: for showing us how to rock the pantsuit.

## Abstract

Diffusion MR Image Processing Tools for Reliable Fiber Tracking Analyses:
Neurosurgery and Radiation Oncology Applications
Kesshi M. Jordan

Neurosurgery and Disconnection Syndrome research have a symbiotic relationship. The human brain is a staggeringly complex system, unique to each individual. Even at birth, there is already incredible diversity to this network, upon which we add a lifetime of experiences, influencing our brain structure and function by the way we use it. One of the best ways to study such a variable and complex system is to see what happens when it is perturbed. Neurosurgical intervention presents a rare opportunity to interact with the human brain in a controlled environment and see what happens when transient or permanent interference occurs. In return, the lessons learned about the relationship between brain structure and function can guide surgical intervention to minimize the risk of surgical injury causing permanent functional deficits. The risk a person is willing to take on to a functional system is a very personal decision; to some people, motor or language function may be what makes life worth living and others are willing to risk deficits to treat a pathology more aggressively. Understanding what damage patterns result in deficits is key to empowering the patient to make these decisions.

The brain's white matter connections can be modeled with Diffusion-Weighted Magnetic Resonance Imaging (DW-MRI) Fiber Tracking (also called tractography), a process by which water diffusion is used to deduce pathways of axon bundles. Neurosurgical applications present particular engineering challenges due to a variety of factors influenced by both the pathology and intervention. This thesis details several tools developed to address these challenges including methods to quality-control tractography streamline datasets, a processing pipeline to model disconnections caused by surgical intervention, a method to translate tractography information to a format tractable for integration with radiation therapy planning, and a pipeline relating electrode stimulation to white matter connectivity. All of the code is open-source so that researchers can use these tools to conduct their own studies.

# Contents

# List of Figures

# Chapter 1

# Introduction

## Background

The relationship between injury to the brain and behavioral changes has been a keystone of human neuroscience research for well over a century [25, 70, 12]. Studying the spatial location of white matter injury, however, is often insufficient to predict functional outcome because disconnection of long myelinated pathways can have widespread effects on systems employing the disrupted circuit, manifesting in apparently inconsistent deficits [70, 71]. Prior to the invention of Diffusion Weighted Imaging (DWI) and, in particular, of Diffusion Tensor Imaging (DTI) in the 1990's, white matter disconnection research in humans was limited to template comparison and post-mortem studies because state-of-the-art imaging was incapable of segmenting white matter fiber bundles for individual subjects in-vivo. DTI enabled non-invasive in-vivo mapping of white matter by inferring the directionality of large, myelinated fiber bundles from the primary direction of anisotropic water movement [11, 108, 133, 130]. DTI and the more complex models that followed [159, 154, 92, 8, 93, 161, 85] have enabled researchers to investigate disconnection syndromes with improved specificity to white matter structure in a variety of clinical applications [106, 105, 171, 42, 74, 118, 90] using fiber tracking, also called tractography, to infer connectivity. Tractography is the process by which the diffusion landscape of the brain, acquired using Diffusion Weighted MRI (DW-MRI), is used to generate a model of the underlying white matter fascicles by tracing out streamlines or other connecting functions [125]. Tractography has been used to

great effect probing the clinical consequence of white matter disconnection by modeling the white matter connectivity on a single-subject level and associating damage to white matter structures with functional deficits [37, 55, 102].

To demonstrate how these models can be used to investigate what connections are threatened by a surgery, an example of an Awake Craniotomy case from UCSF is shown in Figure 1.1. The white matter fascicles associated with language deficits that are thought to be proximal to the tumor were modeled using fiber tracking (For details on the methodology, see [37]). The Arcuate Fascicle (pink), the 2nd and 3rd components of the Superior Longitudinal Fascicle (SLF II&III, also called SLF-ip) (yellow), the Inferior Fronto-Occipital Fascicle (IFOF) (blue), and the Temporal-Parietal component of the Superior Longitudinal Fascicle (SLF-tp) (green) are shown, with a sagittal slice of the pre-surgical image showing the tumor overlaid to the left. It is apparent from the composite view (Figure 1.1 left) that the tumor is very close to several of these fascicle models associated with language function. During the Awake Craniotomy procedure to remove the tumor, the surgeon elicited a word-finding difficulty and semantic paraphasia (a language error in which the patient substitutes a word for the one they intended to say, for example substituting the word "dog" for "cat") [30] when they applied direct electrical stimulation (DES) to the white matter during the resection. This patient came out of surgery with a language deficit and was still non-fluent after several months. However, the patient made a significant recovery somewhere between 4 and 5 months after their surgery.

The goal of this research is to design reliable methodologies for generating these pre-surgical maps and establish how to use them to predict the deficit a surgery is likely to cause on the single-subject level. Clinicians need better tools with which to predict deficits before surgery so that they can counsel their patients and plan interventions to minimize complications and maximize the extent of resection, which will impact patient quality and quantity of life.

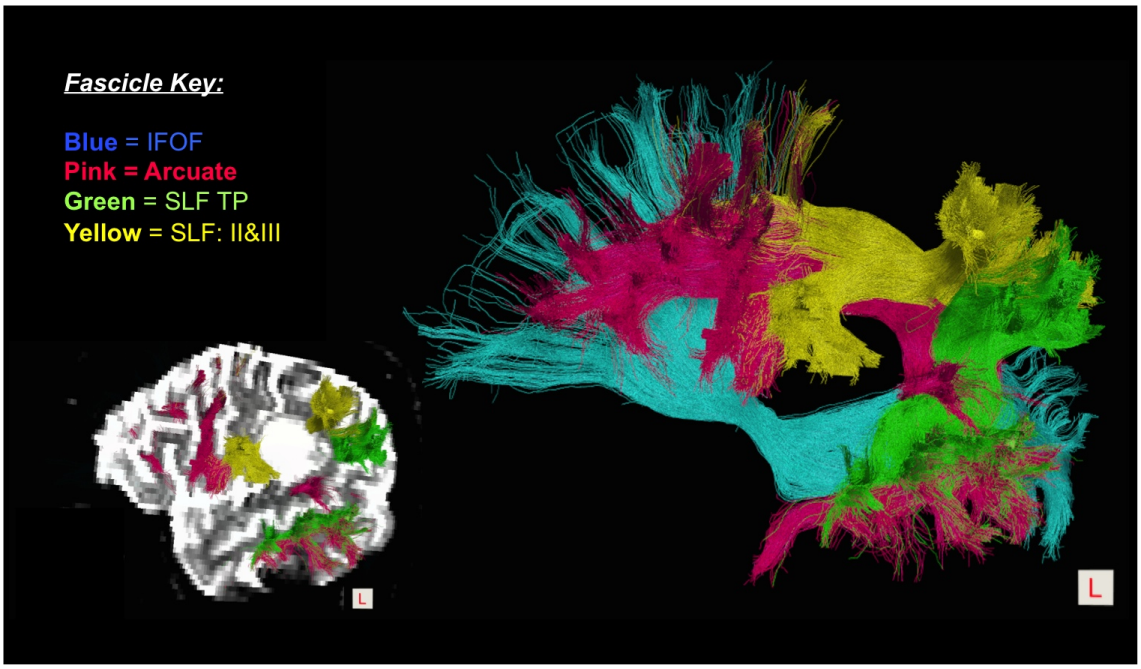


Figure 1.1: Pre-Surgical Rendering of Language-Associated Fascicle Models

*The case shown here is the pre-surgical tractography rendering of a patient's language-associated white matter fascicle models (Right). The fascicle models are shown with the B0 image overlaid to depict the tumor location (circular hyperintensity between the yellow branches).*

## 1.1 Motivation

> *The call to protect life - and not merely life but another's identity; it is perhaps not too much to say another's soul was obvious in its sacredness. Before operating on a patient's brain, I realized, I must first understand his mind: his identity, his values, what makes his life worth living, and what devastation makes it reasonable to let that life end.*

Dr. Paul Kalanithi, When Breath Becomes Air [100]

Brain cancer is one of the top three causes of cancer-related death in people under the age of 40 and the second highest cause of child death in the United States [147]. Surgical resection remains a key aspect of maximizing survival [2], so the importance of learning as much

as possible about the human brain and translating that knowledge to guide neurosurgical intervention on the single-subject level cannot be understated. Between the unique nature of the human brain, the pathology, and the values of the patient, the complexity of this challenge is considerable. The cortex has been mapped extensively, largely because the defining features of the anatomy are readily apparent as soon as the skull is opened and can be easily visualized using non-invasive imaging, like Magnetic Resonance Imaging (MRI). The subcortical bundles running below the surface of the cortex connecting distant and proximal processing components into a vast network, however, are more difficult to visualize and equally (if not more) essential to the preservation of the patient's functional systems.

Diffusion tractography remains the only method of mapping the white matter of the brain non-invasively. As a result, this technology has been increasingly employed as a pre-surgical mapping tool. Tractography methods are highly subject to implementation decisions such as algorithm choice, stopping criteria, initialization parameters, and criteria for defining fascicle volume (e.g. streamline density threshold). At present, pre-surgical methods to isolate a particular fascicle model are highly dependent on a human operator, which introduces additional potential for variability. At UCSF, we have implemented q-ball residual bootstrap probabilistic tractography in pre-surgical planning for motor, optic, and language pathways in brain tumor patients as part of an ongoing research project. The following several sections describe characterization of the variability in our particular tractography implementation at UCSF, described in this chapter. This study investigates inter- and intra- operator variability of fascicle volume definition in a small cohort of control and tumor patient subjects.

## 1.2 Methods

High-Angular Resolution Diffusion Imaging (HARDI) was performed on ten healthy control subjects (3T General Electric Medical Systems Signa Excite) and ten patients with high-grade gliomas both pre- and post-surgery (3T General Electric Medical Systems Discovery MR750) with the following parameters: TR/TE =6425/80 ms, 50 axial slices, 2.2 x 2.2 x 2mm voxel (interpolated in-plane to 1.1 x 1.1 mm), b-value=2000 s/mm2, 55 diffusion gradients, 1 minimally diffusion weighted image (B0). Several language fascicle models

were reconstructed using q-ball residual bootstrap probabilistic fiber-tracking [23] in the left hemisphere by two independent operators, according to the methods reported in [37]. The results were visualized using Trackvis [164]. Language fascicles investigated in all subjects included: the arcuate (AF), the inferior fronto-occipital (IFOF), the superior longitudinal (branches II&III: SLF-ip, and temporal-parietal components: SLF-tp), and the uncinate (UF) fasciculi. The inferior longitudinal (ILF) and middle longitudinal (MdLF) fasciculi were also reconstructed in tumor patients.

For each reconstructed fascicle model, the binary segmentation mask was defined using a streamline density (# streamlines per voxel) threshold. The percent common voxel agreement (PCVA) [116] (Also called Dice Similarity Coefficient) of the binary mask results was defined as the volume of overlap between binary masks divided by the sum of the two mask volumes (scaled by 200) (Equation 1.1). The binary masks defining fascicle volume were based on several thresholds of the original density map (>0, >1, >5, >10, >25, >50, >75 streamlines per voxel). These binary masks were registered to the Montreal Neurological Institute (MNI) space to compare the spatial concordance of intra- and inter-operator fascicle definition across subjects. The intra-rater reliability of the IFOF in controls was also investigated both within a single tracking instance (an operator segmented a fascicle model from the same streamline twice; no variability due to stochasticity of probabilistic method) and between instances of tracking (an operator generated two streamline datasets and segmented a fascicle model from each; variability due to stochasticity of probabilistic method expected).

$$PCVA = \frac{V(L1 \cap L2)}{V(L1) + V(L2)} \times 200 \tag{1.1}$$

## 1.3 Results

Figures 1.2 and 1.3 show the reproducibility of a fascicle volume as a function of streamline density threshold. These results indicate that this probabilistic tractography method tends to have an optimal threshold for maximum PCVA around 5 streamlines/voxel between instances of tracking. This pattern is more prominent in control subjects (Figure 1.2) than in tumor

patients (Figure 1.3). The spatial concordance analysis, which shows the overlap of inter-operator agreement in a cohort of 10 control patients in MNI space, is shown in Figure 1.4 as a function of streamline density threshold and in Figure 1.5 at a fixed streamline density threshold across four different tracks. The blue voxels indicate disagreement between operators, which occurs at the margin of the fascicle model volume. The core of the track has the highest agreement (yellow), with the edges of the fascicle model and the terminations of tracking near the cortex indicating less agreement between operators and/or overlap between patients in MNI space.

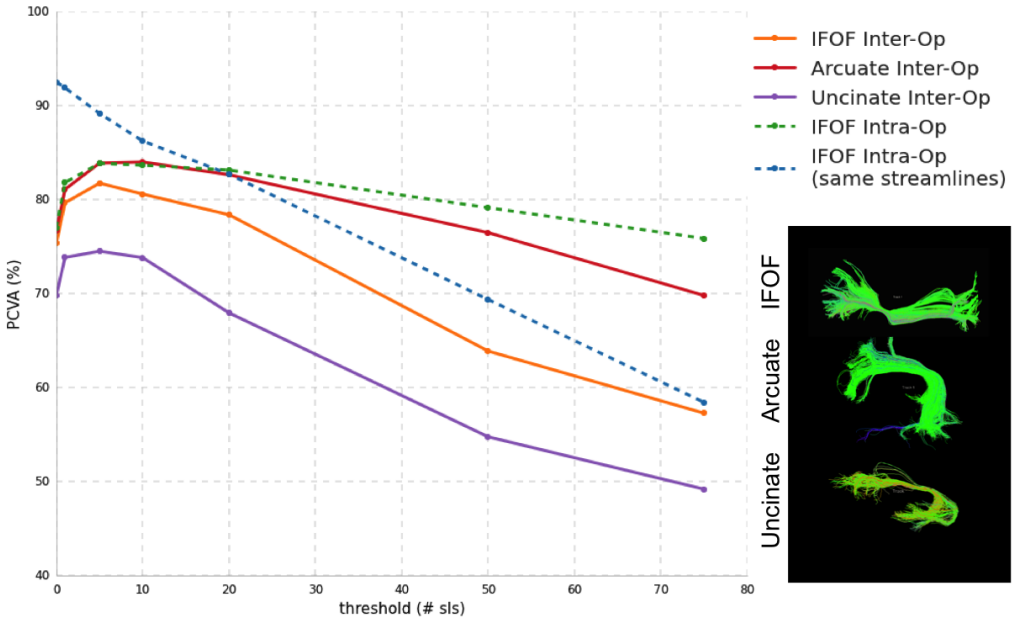


Figure 1.2: Inter- and Intra-Operator PCVA in Healthy Control Cohort

*The chart shows percent overlap between fascicle model volumes inter-operator with unique instances of tracking (solid), intra-operator with unique instances of tracking (green dashed), and intra-operator on a single instance of tracking (blue dashed).*

## 1.4 Discussion

The core of the track generally indicates high agreement (yellow) because it is the trunk of the white matter structure, which should be similar between subjects and consistently

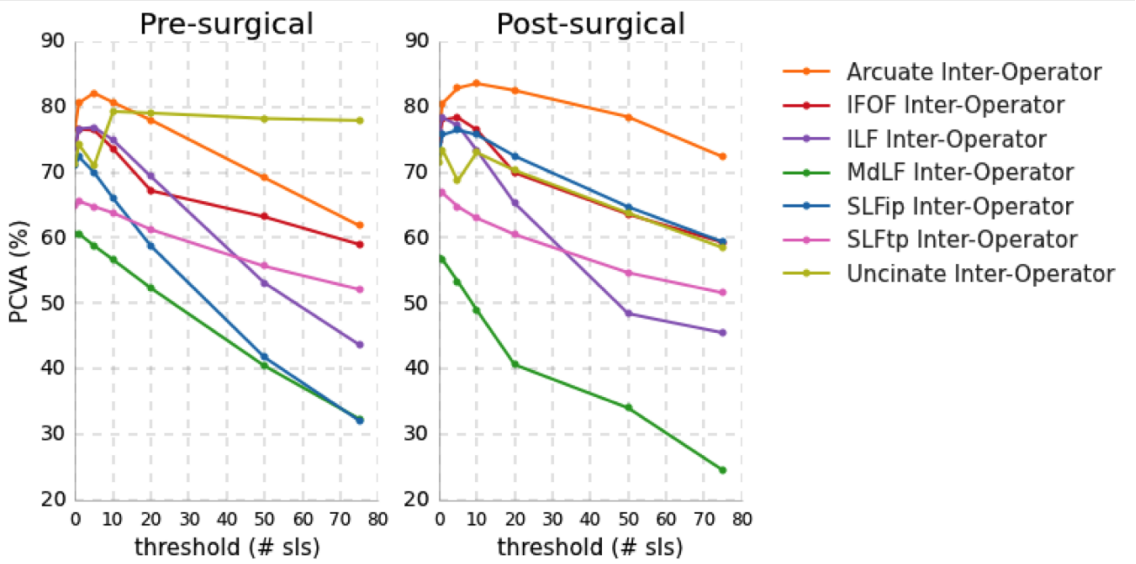


Figure 1.3: Inter-Operator PCVA in Tumor Cohort

*This chart shows percent overlap between fascicle model volumes defined by two different operators with unique instances of tracking in tumor patient subjects pre- and post- surgery*

modeled by tractography methods. The agreement decreases at the periphery of the track due to stochasticity in the probabilistic tracking method, variability between patients, and less consistency between operators. Even post-mortem dissection has limitations establishing terminations of the IFOF in the frontal lobe, where it colocalizes with other bundles like the Arcuate and SLF [121]. Modeling these terminations using lower-resolution water diffusion would not be able to separate uncertainty related to colocalization that challenges dissection, so this performance is not surprising. A threshold of 5 streamlines/voxel largely eliminates disagreement voxels introduced by sparse streamlines at the periphery. Thresholds higher than this cut into the trunk of the fascicle volume, decreasing PCVA. Within a tracking instance (an operator segmented a fascicle model from the same streamline twice; no variability due to stochasticity of probabilistic method), the pure intra-rater reliability is the highest (max 93%) with a threshold of 0, and decays after that (Figure 1.2, blue dashed line) in a manner similar to the between-tracking instance case (an operator generated two streamline datasets and segmented a fascicle model from each; variability due to stochasticity

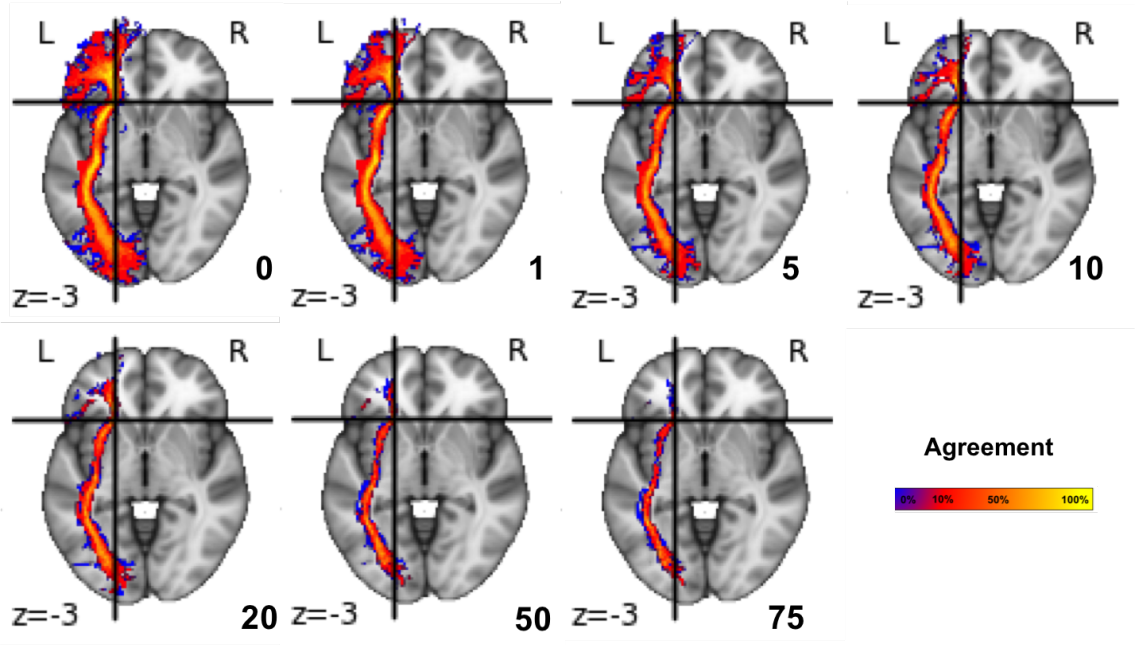


Figure 1.4: Spatial concordance of Inter-Operator Agreement in IFOF Tractography Model: Streamline Density Threshold Dependence

*Spatial concordance of the IFOF Tractography model as a function of streamline density threshold show the expected components of the fascicle at thresholds of 0 and 1. After a threshold of 5, some of the distal components of the IFOF are lost. Significant portions of the IFOF model in the frontal lobe are missing at a threshold of 20 or above, with few cortical voxels being included in the model. (Yellow = high agreement consistently across subjects; Red = low agreement; Blue = disagreement).*

of probabilistic method expected).

## 1.5 Conclusions

Inter- and Intra-Operator reliability shows a marked threshold dependence that varies by fascicle model. These results apply specifically to the methodology presented above. Any changes to the operators, the tractography algorithm, or the data would require this procedure to be reproduced to establish reliability of the new methodology. Methods need to be developed to characterize and minimize these factors for each methodology to ensure that tractography tools are used in a meaningful and reproducible manner.

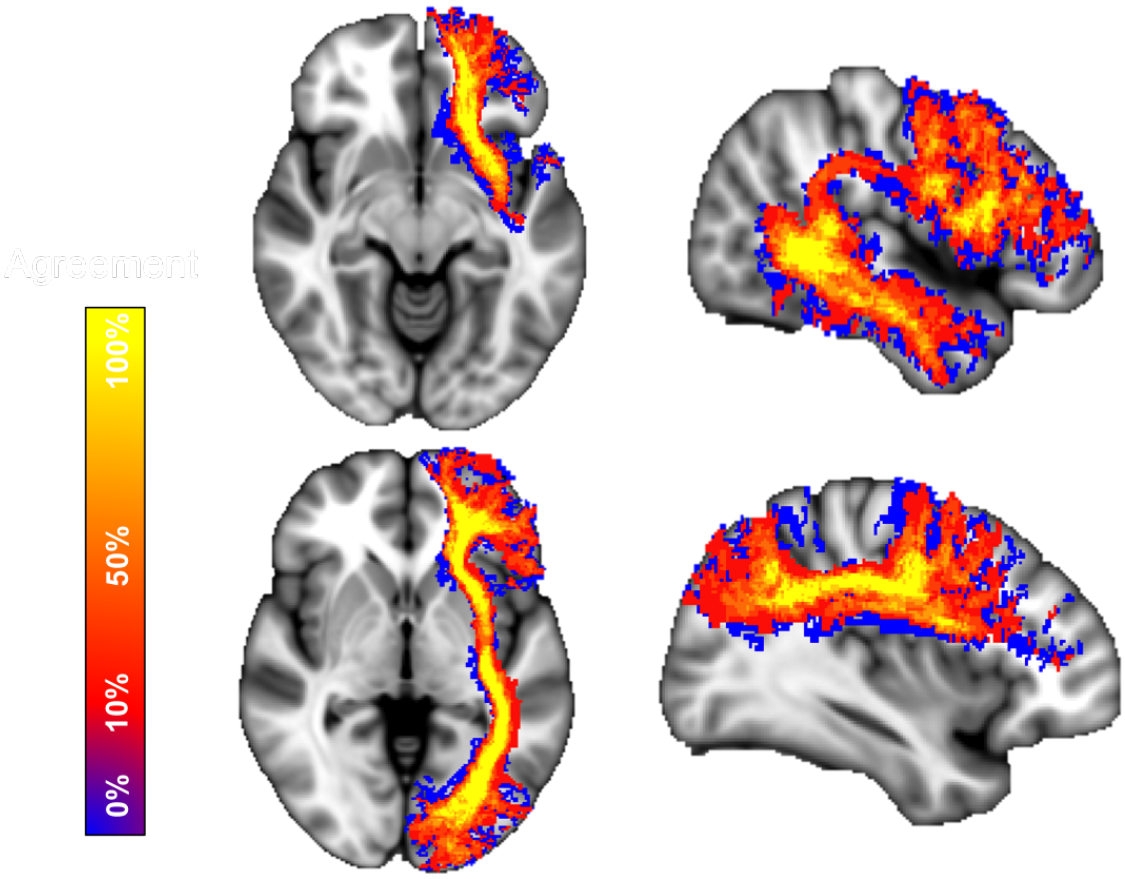


Figure 1.5: Spatial concordance of Inter-Operator Agreement

*Spatial concordance of operator agreement in reconstructions of the Uncinate (top left), Arcuate (top right), IFOF (bottom left), and SLF 2&3 (bottom right) across control subjects using a streamline density threshold of > 0 streamlines/voxel (Yellow = high agreement consistently across subjects; Red = low agreement; Blue = disagreement)*

## 1.6 Author Contributions

This work was conducted in collaboration with Eduardo Caverzasi, Valentina Panara, Anisha Keshavan, and Roland Henry.

## 1.7 Funding

This work was supported by the National Institutes of Health [5R01NS066654-05]. KJ was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

# Chapter 2

# Cluster Confidence Index: A Streamline-wise Pathway Reproducibility Metric

## 2.1 Introduction

Until the invention of diffusion-weighted MRI (DW-MRI) and tractography, the location of white matter pathways could only be inferred using a combination of post-mortem tracing, cortically-based template comparison, and functional imaging. Diffusion tensor imaging (DTI) models water diffusion as a tensor and uses the orientation of the resulting ellipsoid as a surrogate for orientation of the underlying white matter fiber population [11, 133]. The invention of DTI was the first time it was possible to non-invasively map white matter pathways in-vivo, opening up a wealth of possibility for exploring the brain's white matter. Fiber tracking has been used extensively to study a range of fields [168, 43]. Clinically, the use of tractography fascicle models is spreading rapidly in the field of neurosurgery for the purposes of pre-surgical mapping, and as an adjunct to intraoperative mapping. The preservation of white matter structure has been demonstrated as an important factor in preventing permanent functional deficits in neurosurgical patients [57, 89], so tractography techniques are rapidly being deployed in clinical systems.

The diffusion tensor model has well-known limitations, most notably that it is not capable of modeling crossing fiber bundles [97]. Other tractography methods based on more complex models requiring High-Angular Resolution Diffusion Imaging (HARDI) have been developed (i.e. Constrained Spherical Deconvolution, Constant Solid Angle Q-ball) [1, 50, 153, 159, 15], along with complementary probabilistic tractography methods [16, 23, 17] to address the need to adequately model more complex fiber architectures. So-called "higher-order models" used in conjunction with probabilistic algorithms are able to model crossing regions and correspond better with known anatomy [152, 29, 119, 51]. A drawback to upgrading tractography methods to employ probabilistic higher-order models is that they produce noisier streamline outputs that depend more heavily on human intervention to isolate the fascicle model. As the technology migrates from controlled research environments to clinics with varying levels of technical expertise and experience with these methods, we can assume reproducibility in practice will be lower. Translating methods with operator-dependence is difficult to implement uniformly, so developing tools that standardize these procedures as much as possible is a technical prerequisite.

There are many "anatomically constrained tractography" (ACT) methods that outline recipes for the creation of particular fascicle models [33, 37]. These methods typically consist of three steps: a seeding step that places a region-of-interest (ROI) defining the distribution of pathways to be explored, a tracking step to create a streamline dataset, and a targeting step during which a set of inclusion/exclusion ROI's are placed to isolate streamlines that represent pathways characteristic of the fascicle. Exactly how these three steps are executed can vary widely. The seeding strategy typically defines an anatomically-constrained ROI placed on the diffusion colormap (which shows contrast in the white matter) and suggests a "seeding density" to set the number of times the tracking algorithm will be executed in each voxel. The number of seeds placed can depend on a number of factors (the dimensions of the white matter structure, for example). The number of streamlines that reach anatomical targets in the second tracking step also depend on many other factors, such as the pathway architecture, the particular algorithm and seeding strategy used to generate streamlines, the criteria for stopping the tracking (fractional anisotropy threshold or anatomical target, for example), and post-processing. All of these factors can drastically impact

the fascicle model output. Each reconstruction method also introduces artifacts, in addition to underlying imaging artifacts, which can vary with algorithm parameters and underlying anatomy. After tracking, the streamline dataset is targeted using either computationally-derived anatomical targets or those placed by a human operator. Human operators are often used to eliminate low-confidence streamlines in addition to targeting streamlines based on known anatomy, which limits the translation of methods across institutions and introduces operator-dependence that has widely variable results in reproducibility studies [60, 48, 162].

Published ACT methods usually include a seeding and targeting strategy, but do not address the low-frequency streamlines that must be excluded using hand-placed ROIs or the potential variability in how many seeds are needed to fully represent the anatomy-of-interest in a model. The addition of pathology complicates the fascicle model and poses a danger of misrepresenting bundles-of-interest. Probabilistic tractography outputs contain a range of the streamline pathway distribution that can vary considerably subject-to-subject (more-so with pathology). The tractography operator makes many subjective decisions to generate a model of a fascicle that they may tweak in order to create a "reasonable" model that matches their understanding of the anatomy (inflating the seeding density of a sparse model tracking through edema, for example). The operator "cleans" the tractography output by manually placing regions-of-interest (ROIs) to target bundles of streamlines that should or should not be included in the model of the bundle-of-interest. Sometimes these choices are based solely on anatomy, but sparse streamlines are also eliminated because there is lower confidence in the existence of a pathway that was only represented by "outliers". Adding to the complexity, the tractography operator may not know if they have adequately represented this distribution given a unique patient with a unique pathology pattern. Methods like ProbTrackX [16] generate a connectivity distribution with extensive repetition of tractography to address these issues (the default is set at 5000 samples, a sampling density the documentation states should result in convergence), but there is value in visualizing and evaluating the continuous pathway of each individual streamline in the context of its neighbors and any pathology. When using methods that generate a dataset of streamlines (as opposed to a voxel-wise streamline density output, like ProbTrackX), these uncertainties are often addressed by executing a recipe method of seeding/tracking/manual "cleaning" to produce a fascicle model

that is anatomically reasonable (as judged by an expert) and applying a blanket streamline density threshold to eliminate outliers. This can be effective in eliminating outlier streamlines that follow pathways completely alone, but as the streamline density increases the separation between noise and small or difficult-to-track pathways shrinks. There is a need for methods to evaluate the reproducibility of a pathway, instead of a voxel, and to establish at what point the probabilistic distribution of pathways has been adequately represented. Ideally, these methods could be used in conjunction with any fascicle recipe method to objectively generate methodology-specific stable pathway models.

We propose a versatile processing method to address the need for a customizable tool to address outliers uniformly in DW-MRI tractography fascicle modeling. We estimate a Cluster Confidence Index (CCI) using the degree to which streamlines co-localize in bundles, as measured by the minimum average direct-flip (MDF) distance [67]. The idea is as follows: pathways represented by many streamlines following roughly the same trajectory have a higher confidence than those streamlines that follow pathways alone. The approach essentially applies a streamline density threshold to pathways, instead of voxels. In this paper, we demonstrate the results of filtering to remove streamlines representing pathways with low confidence and extend the method to an iterative process in which the fascicle model methodologically converges.

## 2.2 Methods

### Streamline Confidence Metric

This method estimates how reproducible each pathway, represented by a streamline, is with respect to a dataset of streamlines. The Cluster Confidence Index (CCI) is calculated using the minimum average direct-flip distance (MDF) distance, which measures the Euclidean distance between points on subsampled streamlines to quantify similarity between streamline pathways [67]. The CCI for a given streamline is equal to the weighted sum of the MDF metric calculated, pairwise, between the streamline and all of the other streamlines in the dataset (1). The MDF distance implemented in Diffusion Imaging in Python (Dipy) was used [66].

$$CCI_i = \sum_j \frac{1}{MDF(S_i, S_j)^K} \quad \forall \quad MDF(S_i, S_j) < \theta \tag{2.1}$$

The reciprocal of the MDF is raised to the power $\boldsymbol{K}$, which is set by the user. High values of K decrease the weights of streamlines with higher MDF distances, so they will contribute less. Low values of K allow more equal contribution from all MDF levels in the range of contributing streamlines, determined by $\theta$. The maximum MDF distance ($\theta$) that can contribute to the CCI is limited so that huge numbers of distant streamlines cannot inflate the confidence index. Only very similar streamlines with small MDF distances support confidence in the pathway, so we set the default to be 5mm. A simple example of the output of this calculation on a streamline dataset generated from a small ROI in the corpus callosum can be found in Figure 2.1A, with the histogram of CCI values shown in Figure 2.1B.

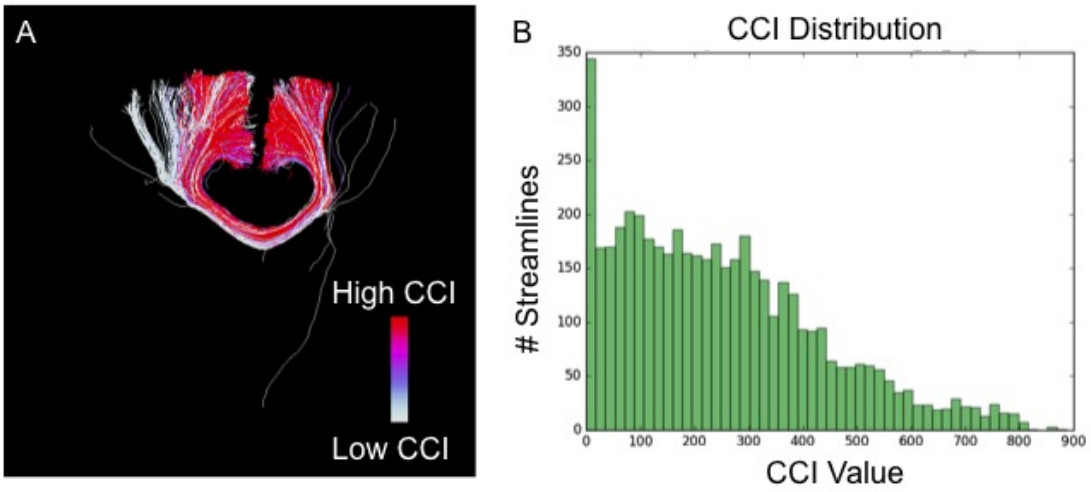


Figure 2.1: Demonstration of Cluster Confidence Index on a Corpus Callosum ROI

***1a*** *Coronal view of a section of corpus callosum streamlines. Streamlines with low support from the surrounding streamlines are colored white. Streamlines with high confidence from the rest of the dataset are colored red.* ***1b*** *A histogram of the CCI distribution in the dataset.*

In applications with sparser streamline datasets, features like fans dominating the pathway, or very long pathways, $\theta$ could be increased to account for a wider distribution of pathways in similar streamlines. In a cohort with a wide range of head sizes, $\theta$ may have to be adjusted for head size. When using this CCI for a new application, it would be beneficial

to try a range of these parameters to get a sense of what settings create the best separation between the bundles-of-interest and outlier streamlines.

## Implemention Details

This streamline-wise confidence estimation framework is agnostic to tractography methodology. For the purposes of this manuscript, the tractography methodologies detailed in this section were employed for demonstration. The code used to generate the CCI and store it in a .trk format that can be filtered interactively using Trackvis [164] [163] is freely available and can be downloaded at [98]. Running the code requires installation of the open-source package Diffusion Imaging in Python (Dipy), along with its dependencies [66].

### Data Acquisition

Two healthy control diffusion datasets were used to demonstrate the method. The first subject underwent an MRI protocol on a 3T General Electric Medical Systems scanner (Discovery MR750) that included a HARDI acquisition with the following parameters: TR 6425 msec, TE 80 msec, axial slices 50, isotropic voxel 2.2 $mm^3$, b value 2000 $s/mm^2$, diffusion gradients 55, minimally diffusion-weighted image 1. This subject was used for all figures, except those pertaining to the iterative tracking demonstration. This study was carried out in accordance with the recommendations of the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by our institution, with written informed consent from all subjects. The protocol was approved by the Committee on Human Research at the University of California, San Francisco.

A publicly available healthy control HARDI dataset was used as the second subject [138] to demonstrate the iterative tracking procedure. This dataset has the following parameters: axial slices 76, isotropic voxel $2mm^3$, b value 2000 $s/mm^2$, diffusion gradients 150, minimally diffusion-weighted images 10.

### Fiber Tracking

Unless otherwise stated, Q-ball Residual Bootstrap Fiber Tracking was performed using the open-source software package Diffusion Imaging in Python (Dipy) [66, 22] to gener-

ate streamlines for each demonstration (excepting the iterative tracking example). During residual-bootstrap tractography [23], principal fiber orientation was estimated at each step by computing a bootstrapped orientation distribution function (ODF) and identifying the peaks. The constant solid angle variant of the ODF function described by Tristan-Vega et. al. [157, 155] was used. ODFs were fit using even order spherical harmonic functions up to order 4. ODF peaks less than 45deg from a larger peak and small peaks with values less than one-fourth of the maximum of the ODF were excluded. The principal fiber orientations from the bootstrap ODFs provided the distribution of fiber tracking directions. Tracking was terminated upon meeting either of the following criteria: FA threshold of 0.15 or maximum turning angle of 60deg [29]. Results were visualized using Trackvis [164].

### Fascicle Modeling

Several "recipes" for fascicle modeling were used to demonstrate realistic applications of this streamline-wise confidence metric. Fascicle models of the Arcuate Fasciculus (AF), the 2nd and 3rd components of the Superior Longitudinal Fasciculus (SLF2&3), the temporo-parietal component of the Superior Longitudinal Fasciculus (SLF-tp), the Inferior Fronto-Occipital Fasciculus (IFOF), and the Uncinate Fasciculus (UF) were built by seeding ROI's defined on a fractional anisotropy (FA) colormap according to the methods described in [37] at a seeding density of $7^3$ seeds per voxel. In short: the streamline dataset from which the IFOF and UF were segmented was created by seeding a cross-sectional ROI across the external/extreme capsule on a single coronal slice at the level of the anterior commissure. The streamline dataset from which the AF and SLF2&3 were segmented was created by seeding a cross-sectional ROI across the SLF, on a single coronal slice at the level of the isthmus of the corpus callosum. The streamline dataset from which the SLF-tp was segmented was created by seeding the inferior parietal lobule on a transverse ROI covering the supramarginal and angular gyri.

## Iterative Tracking

For the demonstration of iteratively applying this framework to establish an appropriate seeding density, a small 10x10x1 voxel square ROI of voxels with orientation distribution

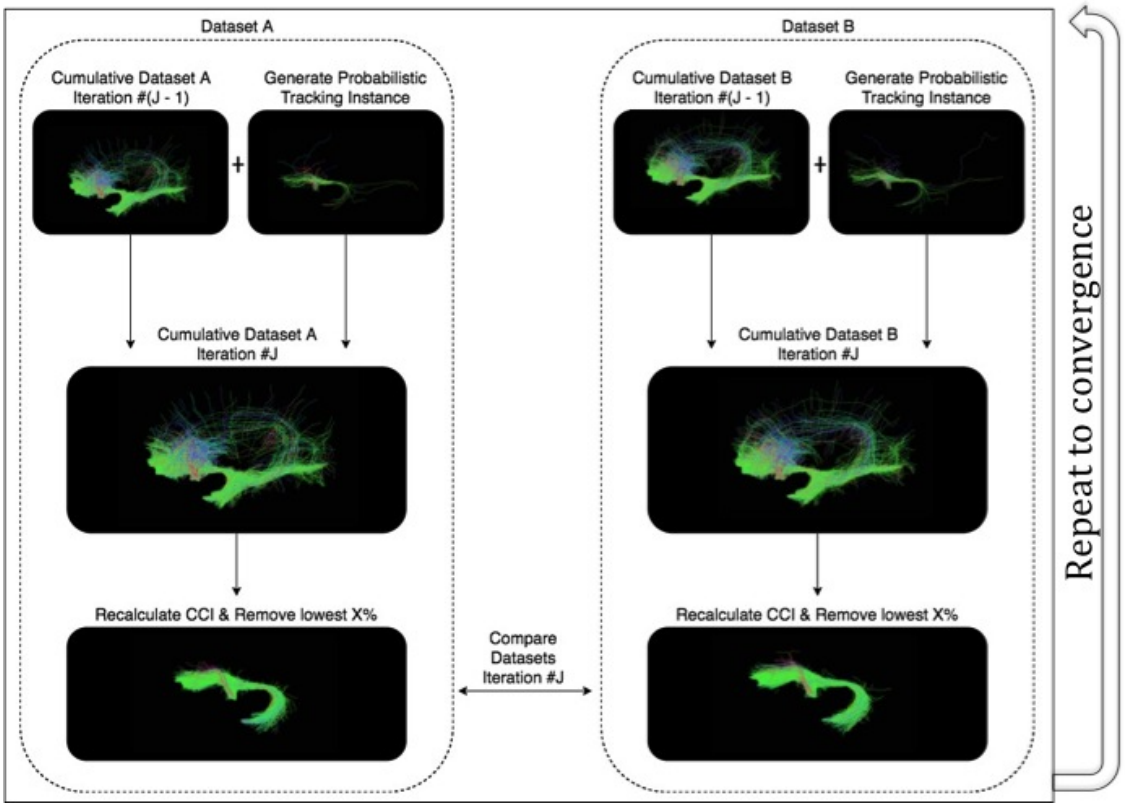


Figure 2.2: Flowchart to generate two equivalent probabilistic tractography datasets in parallel using the Cluster Confidence Index (CCI) to filter low-confidence streamlines

*For each iteration (J), tracking is executed once with the established seeding/tracking scheme and the streamlines added to the cumulative dataset (from iteration J-1). The CCI is then calculated on this cumulative streamline dataset and the lowest X% discarded. This procedure is performed twice in parallel and, after each iteration, the streamline datasets are compared to evaluate the "methodological reproducibility" of that iteration.*

functions (ODFs) indicating fiber crossings was selected on a transverse slice of the frontal lobe to demonstrate convergence of tractography in a region with complex anatomy. A flowchart of this procedure can be found in Figure 2.2. The tracking algorithm/parameters were applied iteratively twice in parallel (Datasets A and B). For each iteration (J), a single instance of tracking was performed with the seeding density=1x1x1 using the method detailed in the Dipy software package CSA tutorial [4]. This output was added to the cumulative dataset from iteration (J-1). The Cluster Confidence Index was calculated for this new cumulative dataset (J) with the default parameters (theta=5, k=1), and the dataset

was thresholded at the Xth percentile (60th in this demonstration); in other words, the X% streamlines with the lowest confidence were discarded from the dataset. This procedure was performed twice in parallel and, after each iteration, the streamline datasets are compared to evaluate the "methodological reproducibility" of that iteration. Two comparison methods were used: (1) the percent volume overlap (PCVA), as described in Equation (2) and [116]; and (2) the fopt optimization value for registration of the two streamline bundles using the streamline-based linear registration method by [68].

$$PCVA = \frac{V(L1 \cap L2}{V(L1) + V(L2)} \times 200 \tag{2.2}$$

## 2.3 Results

### Cluster Confidence Index

The Cluster Confidence Index (CCI) is demonstrated in Figure 2.1 on a sagittal ROI in the corpus callosum, with a histogram of the CCI distribution for this streamline dataset. Several outlier streamlines are apparent in this dataset; these outliers travel a pathway alone, deviating from the expected route, so they have very low confidence (shown in white). Most of the streamlines follow the characteristic pathway of the corpus callosum in coherent bundles, forming the shape we expect from anatomical knowledge (shown in red). This view can be used to evaluate the performance of the CCI. The parameters discussed in Equation (1) (K and theta) can be adjusted to maximize the separation between desirable and outlier streamlines, the seeding density can be inflated to better describe the distribution of pathways passing through the seeding ROI, and the length filter adjusted to focus on a specific scale of bundles. Figure 2.3 shows several tractography outputs used routinely at UCSF to model fascicles for pre-surgical planning and an example of whole-brain seeding. Each segmentation application has unique challenges in terms of distribution of pathways that make up the bundle-of-interest and those that must be eliminated. The IFOF and UF (Figure 2.3 A) are easy to separate; the UF hooks inferior/anteriorly into the anterior temporal lobe (down and to the left on this sagittal view), and the IFOF continues posteriorly, to the occipital lobe (to the right on this sagittal view). The Arcuate and SLF2&3 (Figure

2.2 B) are more difficult to separate and have many noise streamlines that overlap with portions of both fascicle models. The CCI is an additional tool that can be used to separate desirable streamlines from undesirable.

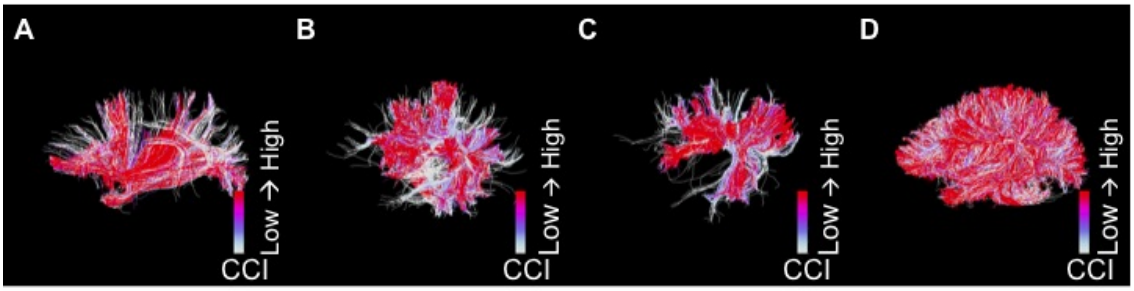


Figure 2.3: Fascicle Model Examples of Cluster-Confidence Index (CCI)

*Streamlines datasets are viewed from the left side colored from low confidence (white) to high confidence (red).* ***A*** *External/Extreme Capsule seed used to segment Inferior Fronto-Occipital Fasciculus and Uncinate Fasciculus;* ***B*** *SLF seed used to segment Arcuate Fasciculus and SLF II & III;* ***C*** *Inferior parietal Lobule seed used to segment temporal-parietal component of the SLF;* ***D*** *Whole brain tractography.*

## CCI As an Interactive Tool

The code provided produces an alternative .trk file with the CCI embedded in the structure. This file can be opened in Trackvis and filtered by CCI interactively, as shown in Figure 2.4.

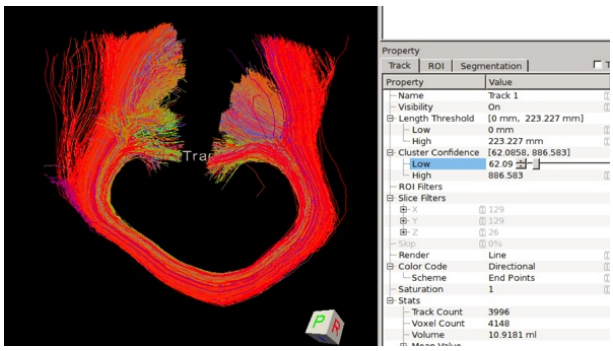


Figure 2.4: Interactive Trackvis Implementation

*Streamlines can be filtered based on the CCI interactively using Trackvis.*

The interactive CCI slider bar is helpful in the manual segmentation of fascicle models. The initial tractography output can have many outlier streamlines obscuring the structure of the underlying fascicles that need to be targeted, especially in cases with pathology like a tumor. The user can temporarily filter out low-confidence streamlines to place targets, and then proceed with segmenting the fascicle model. Also, the CCI can uniformly and easily filter out all of the outlier streamlines when the model is finished so that the user does not need to waste time targeting sparse unwanted streamlines, individually. This process is demonstrated in Figure 2.3 with the tractography output used to segment the Arcuate and SLF2&3 fascicle models. The initial tractography output (Figure 2.5 A) has many streamlines obscuring the view of the fascicles we are trying to segment. As the CCI minimum threshold is increased (Figure 2.5 B-D), the outlier streamlines disappear and it is easy to see exactly where the targets need to be placed to isolate the Arcuate and SLF2&3 fascicle models. A user should be careful when applying this tool as an outlier removal threshold to any dataset with a wide range of streamline trajectories because there may be overlap between desirable and undesirable pathways.

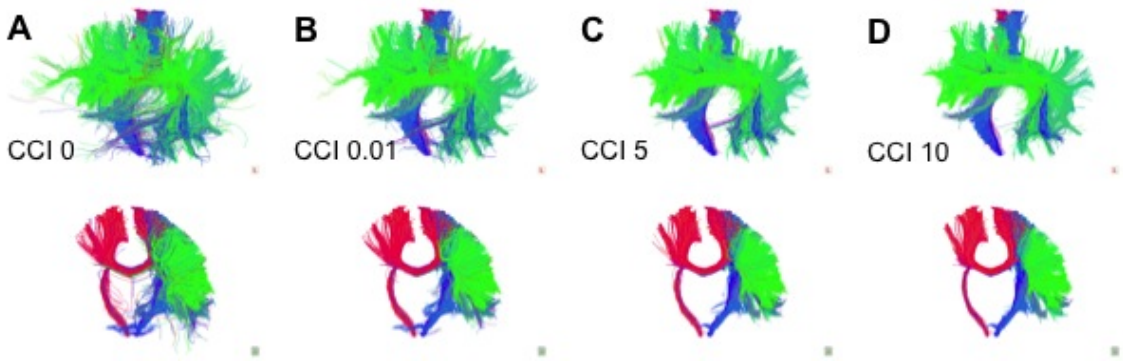


Figure 2.5: Superior Longitudinal Fascicle Complex Model Filtered by CCI

*Interactive fascicle cleaning using the Cluster-Confidence Index on the fascicle output from which the Arcuate and SLF2&3 fascicle models are segmented:* ***A*** *The original set of streamlines an operator is presented with when they seed a conservative cross-section of the left SLF viewed sagittally through the patient's left ear (top) and coronally looking at their face (bottom);* ***B-D*** *The streamlines filtered at increasing levels of Cluster-Confidence Index (CCI) from 0.01* ***B*** *to 10* ***D****.*

Using the CCI interactively is useful for exploring a noisy tractography output to quickly

get an impression of the dominant streamline bundle architecture in a dataset. For example, in a whole-brain tractography dataset (Figure 2.6), increasing the CCI filter to 10 reveals the underlying structure of the SLF/Arcuate complex, IFOF, and UF. This approach can be used to explore the tractography around a noisy pathology, like a tumor, to get an impression of how the pathology is affecting the tractography algorithm. This knowledge can help to interpret fascicle models built for the patient, suggesting how much confidence we have in peri-tumoral streamlines, and facilitate efficient manual targeting.

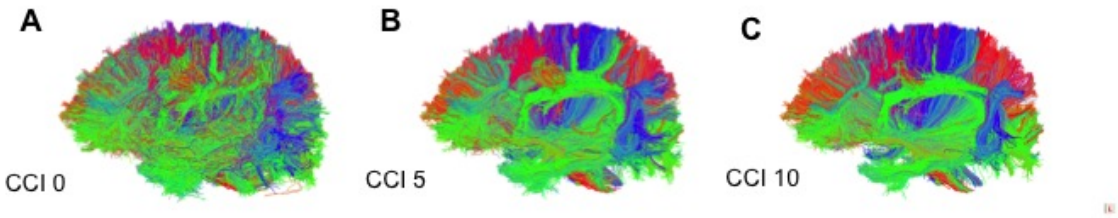


Figure 2.6: Whole Brain Dataset Filtered by CCI

*The raw set of streamlines an operator is presented with when they seed all of the white matter in the brain viewed sagittally through the patients left ear is noisy and difficult to work with (**A**). The streamlines are filtered at increasing levels of Cluster-Confidence Index (CCI) from 5 (**B**) to 10 (**C**) to reveal the major bundle architecture of the dataset.*

## CCI As a Tool for Seeding Density Selection

Standardizing seeding density is not a guarantee for creating comparable fascicle models across subjects, especially in pathological cases. Under-seeding risks an incomplete model of the distribution of pathways present in the white matter. It is easier to identify this problem in controls than it is in patients with pathology because confidence in the fascicle model is based on known anatomy (established by dissection), which may not be true in the presence of pathology. The iterative approach described in this manuscript (Figure 2.2) estimates the methodological reproducibility by conducting two instances of iterative tracking in parallel and comparing the resulting bundles. In Figure 2.7 we see that the results of the parallel processing streams look similar (top vs. bottom), but fill out as the number of iterations increases (left to right). If we plot the two similarity metrics, percent volume overlap

(PCVA) (Equation 2.2) and the optimization parameter from Streamline Linear Registration (fopt) [68], as the iteration increases from 1 to 100 (Figure 2.8 top), both of them converge. PCVA converges to just over 80% overlap, and the fopt streamline bundle difference from the Streamline Linear Registration optimization parameter (fopt) drops off to around 5 within the first 20 iterations, then continues to decrease slowly. Plotting the difference in PCVA and fopt between iterations (Figure 2.8 bottom), we can see that the metrics change minimally after the initial 20 iterations and a tolerance could be set to terminate the procedure. At the point where the difference between these parallel processing streams converges, we assume that the distribution of pathways present in the seeding ROI has been adequately represented because it is reproducible and not changing with the addition of streamlines. Removing the lowest confidence streamlines helps to smooth this convergence by eliminating outlier streamlines. The choice of removing the lowest 60% of CCI streamlines emphasizes the length bias (toward shorter streamlines) to model more local connectivity (Figure 2.9). A tolerance and CCI threshold could be chosen to standardize seeding across patients (instead of seeding density) based on this convergence to ensure that the distribution of pathways has been explored equivalently. Figure 2.10 shows a comparison between the volume measurement of the fascicle models at each step of iterative tracking generated using a density map (# streamlines per voxel) and/or CCI threshold. The comparison was made using both the PCVA (Figure 2.10 A) and the fascicle model volume (# voxels) (Figure 2.10 B). The raw streamline output (dark blue) performs the worst, in terms of reproducibility; the PCVA is very low and there is about a 5% difference in volume between two outputs from two identical tracking instances after 100 iterations. A high PCVA can be achieved using voxelwise thresholds of 5-10 streamline/voxel (red and purple) or a CCI threshold of the 60th percentile, but the volume is more reproducible with the CCI approach (yellow and light blue).

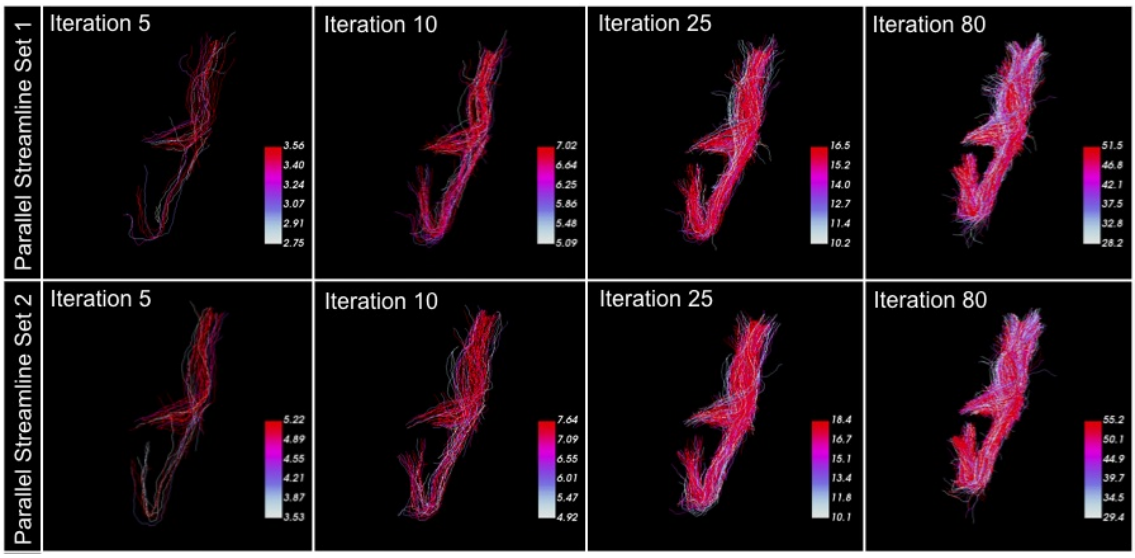


Figure 2.7: Iterative Tracking with CCI Filter

*Tracking is iterated in parallel to create two sets of streamlines. At each iteration, low-confidence streamlines are removed and the streamlines generated from that tracking iteration are added to the aggregate dataset. Since this is occurring twice in parallel, methodological noise can be estimated by comparing the equivalent volumes produced by binarizing the two streamline sets.*

## 2.4 Discussion

### Uniform Elimination of Outliers

The choice to uniformly eliminate outliers from a streamline dataset requires an investment in methodology development, but has many advantages over manual selection and setting a streamline density threshold. Manual selection of outliers cannot be reported rigorously in a manuscript and is labor-intensive. Thresholding based on the CCI is intended to replace the use of density maps, which are commonly employed to increase confidence in the volume segmented as a fascicle model. The use of CCI-based thresholds is advantageous over streamline density-based thresholds because the CCI works in the space of pathways, rather than voxel space. As an example: a density map will be artificially inflated in regions of crossing fibers because the confidence metric "# streamlines per voxel" includes a count from multiple fiber populations. The CCI assigns each individual streamline a confidence

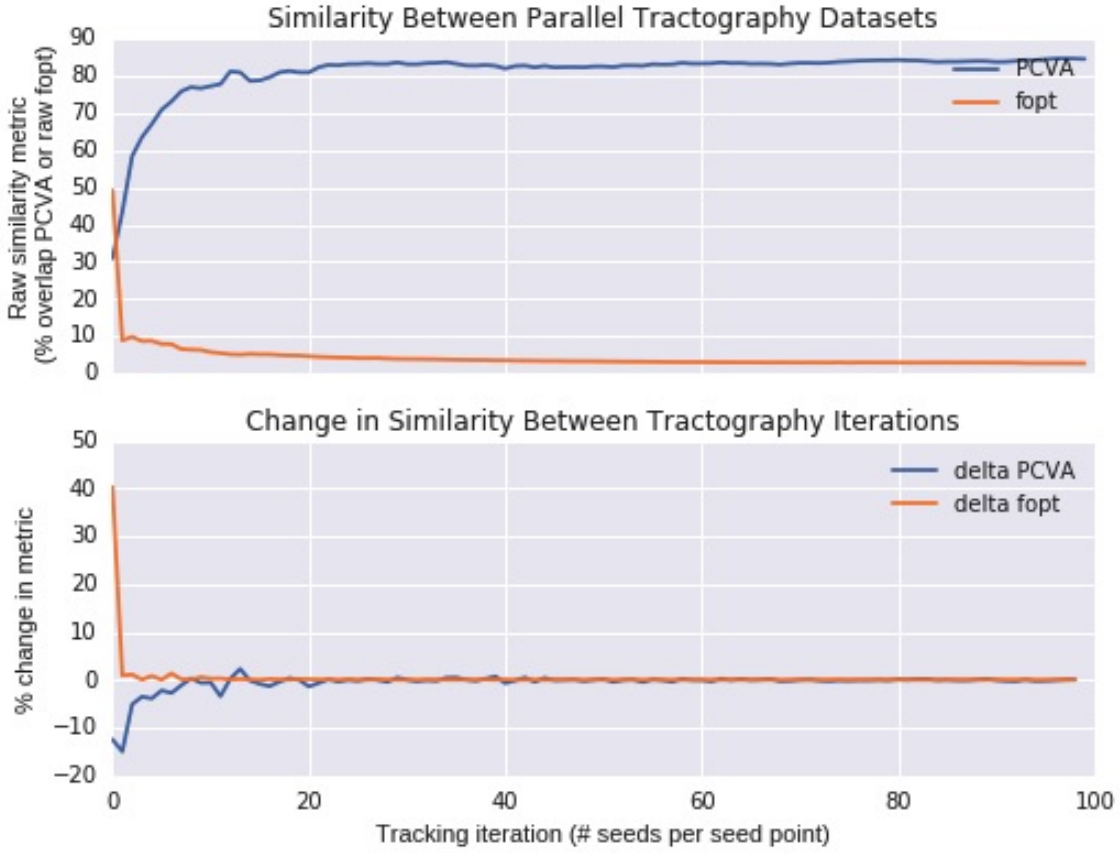


Figure 2.8: Iterative Tracking with CCI Filter Convergence

*Both measurements of similarity between independent streamline datsets (PCVA = percent volume overlap and fopt = bundle similarity [68]) change between tracking iterations less and less as streamlines are added to the cumulative dataset, converging toward zero.*

level based on agreement of local streamlines. It is not surprising that the volume generated by thresholding based on pathway density (using the CCI) is more reproducible than that generated based on voxelwise density (using a streamline density map), as is shown in Figure 2.10.

Other methods have been developed to uniformly eliminate outliers, but they tend to be specific to the applications for which they were designed. Several approaches have employed the Quickbundles algorithm [36, 44] (which uses the MDF distance to cluster streamlines) to build methods for dealing with "messy" streamline datasets because it is very fast and a good building-block with which to simplify complex streamline datasets. [44] developed a very

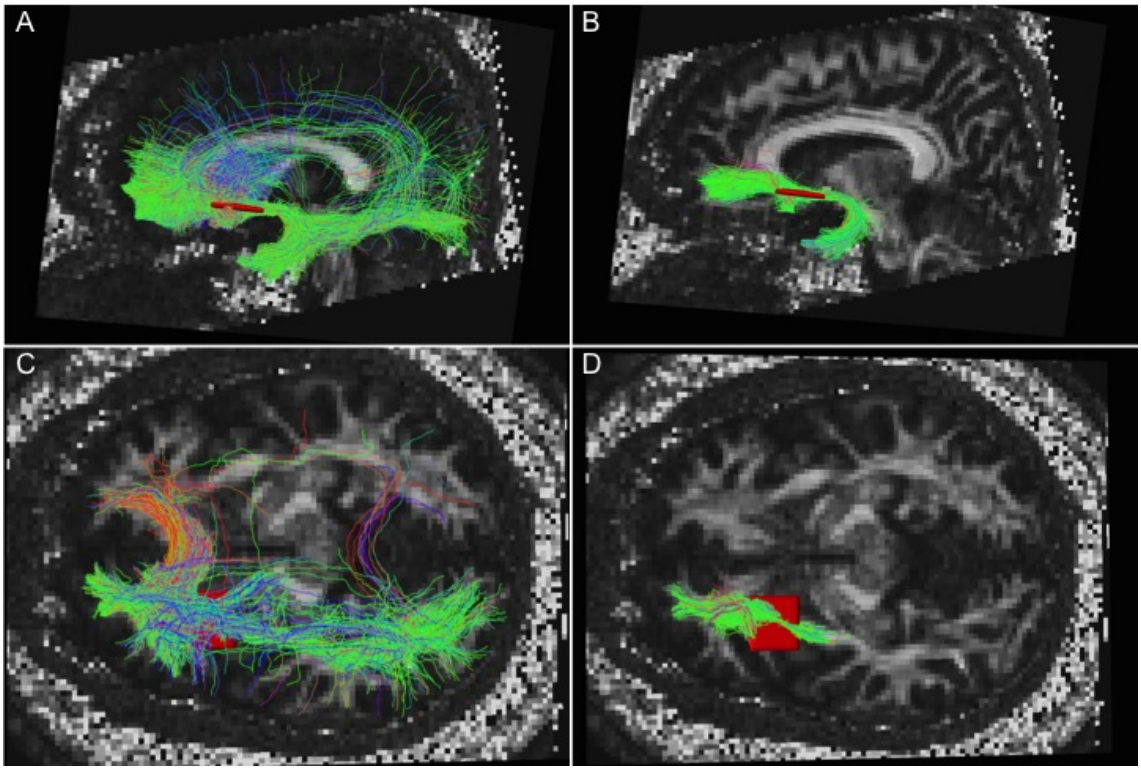


Figure 2.9: Iterative Tracking with CCI Filter Streamline Output

*Panels **A** and **C** show the raw streamline dataset after 80 iterations. Panels **B** and **D** show the same dataset thresholded at the 60th percentile CCI to select for the highly confident local connectivity of the tracking ROI (red). There is an obvious length bias to this high CCI threshold, but the approach shows two anatomically-defined connections quite strongly and cleanly: the Uncinate fascicle hooking down into the temporal lobe, and the Fronto-Orbitopolar Tract [35]. To reconstruct the inferior fronto-occipital fascicle (IFOF), an occipital lobe target would have to be integrated with tracking so that the length bias does not influence the final result.*

similar approach to the one presented here using the Quickbundles algorithm hierarchically to identify outlier streamlines. The method presented in this article is a variation on the same concept, but this implementation uses the MDF distance to evaluate each single streamline for how much support it has from the surrounding streamlines in the dataset. This approach is more general with several parameters that can be tuned for a specific application to optimize performance that may vary with respect to scale of anatomy being studied, data quality, seeding approach, pathology, or target use of the fascicle model.

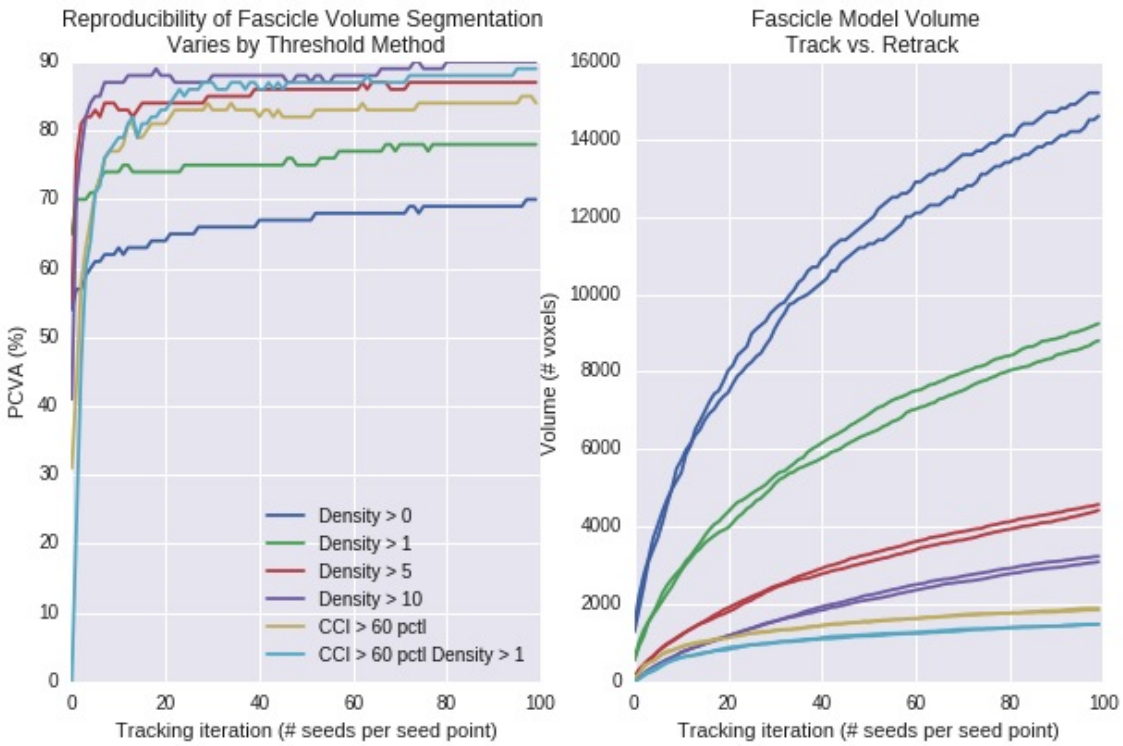


Figure 2.10: Iterative Tracking with CCI Filter Reproducibility

***A**: As the seeding density is increased, an 80% or greater percent overlap (PCVA) can be achieved either with a density threshold of 5-10 or CCI threshold at the 60th percentile. **B**: Thresholds based on the 60th pctl CCI maintain a consistent volume more so than those based on density maps, indicating a robust track-retrack methodological reproducibility.*

The method presented here is a helpful tool both in streamlining and standardizing manual segmentation of tractography outputs to create fascicle models. Applying even a very low threshold to eliminate the obvious outliers improves standardization and translation of methods. Fascicle models shown in figures often involve some sort of outlier cleanup that cannot be fully described in the methods ("outliers were removed manually"). Using a low CCI threshold makes the figure presentable in a manuscript by eliminating the outliers obscuring the result, but the action can be accounted for in the methods and reproduced.

The interactive filtering capability that can be used with the Trackvis viewer makes manual streamline "cleaning" much less labor-intensive; it enables a user to quickly evaluate the reproducible structure of a tractography output, and to eliminate outliers uniformly

with a slider-bar, instead of targeting them manually. Calculating the CCI for a dataset scales N * N * P (N being # streamlines and P being # points per streamline), so it is advantageous to minimize the number of streamlines or points per streamline. The CCI can be recalculated at multiple stages as the user refines the bundle, if needed, but large datasets may need to be reduced to limit calculation time. This could be done by calculating CCI with less datapoints or limiting the calculation to bundle centroids until the dataset has been refined to the point that calculation time is acceptable to the user. The method works on the streamlines, so the user is not restricted to any particular tractography method, and parameters can be tuned to optimize use for a given application.

By applying these methods iteratively, as demonstrated in Figures 2.7-2.10, the user can ensure that only methodologically reproducible streamlines are considered for inclusion in a model. This also protects against the risk of under-seeding a structure; if all fascicle models are tracked iteratively until convergence, as demonstrated in Figure 2.8, the user can be confident that they have represented the distribution of pathways in the seeding region equivalently across differing anatomy or pathology.

## Considerations

It should be noted that there is a length bias to the CCI; as the CCI minimum threshold is increased, many very short streamlines are shown to have the highest confidence. We argue that this is reasonable given the nature of tractography. Errors accumulate with distance so we do, in fact, have more confidence in the pathway of short streamlines than we do in longer streamlines that had the opportunity to accumulate more error as they traversed a further distance. In Figure 2.9 A/C, the raw streamline output includes bundles from the Corpus Callosum, the Inferior Fronto-Occipital (IFOF), the Uncinate (UF), and Frontal Orbitopolar (FOP) [35] fascicle models. After a high CCI threshold is placed on the dataset, only the UF and FOP remain (Figure 2.9 B/D). A length weight would have to be incorporated to set up an iterative tracking procedure for longer fascicle models, like the IFOF, or the base set of streamlines from which the CCI is calculated controlled. Incorporating an appropriate minimum length filter is a good idea to avoid the inflation of high-confidence short streamlines and make the streamlines that are being considered for inclusion the focus of the CCI

distribution to aid in separating the desirable from undesirable streamlines with a single threshold (there is an option to do this in the provided code). When using the CCI as a tool to aid in elimination of outliers from a particular fascicle model, it may be helpful to re-calculate the CCI after targeting the fascicle model to isolate the bundle(s)-of-interest. The CCI calculation could be integrated with a viewer to enable a user to periodically calculate the CCI as they narrow down the selection of streamline bundles.

Parameter exploration/tuning should be explored when tracking a new structure or cohort. Theta sets a hard limit on how local streamlines must be to contribute to the CCI, at all. The K parameter determines how steep the distance dropoff is of the weight each streamline "vote" gets toward the CCI. For example, if streamline A follows exactly the same path as its neighbor, it suggests more support for that particular pathway than another streamline (streamline B) that diverged at the halfway point. The K parameter determines how much more weight streamline A's vote has than streamline B's; Theta decides if streamline B gets a vote, at all. It should be noted that setting these parameters across subjects poses a potential problem; because of the length bias in confidence, equivalent streamlines for people with larger head sizes will have lower confidence because their longer streamlines have more distance over which to accumulate errors. Further work is needed to incorporate dynamic weighting of distances based on some head-size-related metric.

Ideally, tractography should be performed to a convergent result in a standardized manner to maximize reproducibility. The scale of the CCI depends heavily on seeding strategies and the original set of streamlines. Further work is needed to optimize seeding strategy for comparison across timepoints and patients. The iterative demonstration, shown above, is one way to evaluate when the system has been tracked enough to represent the complexity of the underlying structure, but further work is needed to develop standards to force the convergence of this system. One potential method would be to predict the diffusion signal back filtered streamlines using an approach like the Linear Fascicle Evaluation [132] approach to determine convergence.

## Why do we need these types of tools?

A study by [60], evaluated the anatomical accuracy of fornix tractography across many algorithms using a subjective rating system (1:best to 4:worst) based on anatomical accuracy and number of streamlines outside of anatomical boundaries. The accuracy scores ranged from 1.1 to 3.6 and the incorrect fiber scores ranged from 1.6 to 3.5. The authors' evaluations of anatomical accuracy and fibers outside of anatomical boundaries can be considered surrogates for sensitivity and specificity, thus subject to the stereotypical trade off characterized by an ROC curve. The most accurate approach (accuracy 1.1) suffered from false positives (incorrect fibers 2.4) and the approach with the least fibers outside of anatomical boundaries (incorrect fibers 1.6) did not perform well in accuracy (accuracy 2.1). Feigl et al. stressed that the manual intervention commonly employed to "clean up" false positives with placement of exclusion ROIs creates a danger for excluding displaced fibers that represent genuinely displaced white matter bundles [60].

The CCI tool, presented here, does not eliminate the subjectivity problem that Feigl et al. was describing, but it at least establishes a standard for uniformly eliminating them and reporting exactly how that was done. Manual segmentations of fascicle models often require some outlier cleanup that cannot be detailed in the methods because it has an element of randomness to it. If the operator used a method like CCI thresholding to eliminate outliers, the parameters of the tracking and CCI calculation/filtering should provide some grounds from reproducibility. Sparse streamlines are sometimes random, but not always. Some sparse streamlines follow small or difficult-to-track structures that would have been more robustly represented with a higher seeding density. Seeding density requirements can vary by patient; sometimes a pathology siphons off streamlines so that the underlying anatomy is not the dominant pathway represented in the streamline dataset, or edema causes streamlines to meet a stopping condition early at a higher frequency, so more seeds are needed to represent the underlying pathway. Applying an iterative tracking procedure, like the one described in this manuscript, ensures that the outliers removed are not a symptom of under-seeding. The only manual interventions should be anatomically-defined ROIs for targeting the desired bundle-of-interest and a discrete number of well-defined artifacts that can be described in the methods. Minimizing the extent of manual intervention and standardizing the necessary

components are crucial steps in moving forward with the ongoing clinical translation of tractography methods.

Reconstruction of fascicles using a tensor model (DTI) is easier to execute because far less false-positive pathways are represented. Methods that are sensitive enough to represent crossing regions in the brain (Probabilistic Q-ball, Probabilistic Constrained Spherical Deconvolution, etc.) are also prone to outlier streamlines, thus have a tradeoff of low specificity and require "clean up" of false positives by manual intervention. By limiting the number of low-confidence pathways a tractography operator is presented with and minimizing the subjective decisions made, with respect to outliers, transitioning to higher-order models becomes less intimidating, in spite of their high level of false positives.

New tractography methods are being developed constantly, and the variety of implementations used across the field is already considerable. Beyond the question of DTI vs. HARDI methods, the addition of multiple b-values ("shells") opens up the possibilities to many more models that can even better characterize underlying white matter structure [8, 94]. Versatile tools to facilitate uniform quality control of tractography outputs are needed to support the wide variety of methods used to build fascicle models.

## 2.5 Conclusion

The Cluster Confidence Index method presented in this manuscript objectively calculates the confidence of a particular pathway in a dataset of streamlines by comparing it to the distribution of streamlines following similar pathways. This method can be used to uniformly remove outlier streamlines by placing a low CCI minimum threshold on all streamline datasets, interactively to expedite manual cleaning of fascicles, or iteratively to establish a reproducible tractography methodology. This method provides a framework for reporting exactly how outliers were eliminated so that methods can be more accurately presented in manuscripts. The approach is tractography algorithm-agnostic; it will work with an streamline-like dataset. The implementation provided works on a standard .trk file and can be filtered interactively by CCI in the Trackvis viewer. The code is free and can be downloaded from [98].

## 2.6 Author Contributions

This work was conducted in collaboration with Bagrat Amirbekian, Anisha Keshavan, and Roland Henry.

## 2.7 Funding

This work was supported by the National Institutes of Health [5R01NS066654-05]. KJ was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

## 2.8 Acknowledgments

The authors would like to thank the Neuroimaging in Python (NIPY) development community, all of our research subjects, and the providers of the publicly available MRI dataset used in this research.

# Chapter 3

# Cluster-viz: A Tractography QC Tool

## 3.1 Introduction

When tractography algorithms are used to create an anatomically constrained model of a fascicle, the output of the processing can contain many streamlines that are not part of the bundle-of-interest. Using methods that leverage High Angular Resolution Diffusion Imaging (HARDI) datasets by employing models like Constrained Spherical Deconvolution [153, 154] or Q-ball [159, 160, 23] increases the sensitivity of the method (compared to the simpler tensor model), but generates many more streamlines that must be excluded. Automatic classification methods have been developed [172, 169], but pathologies (e.g. tumors) present in patient populations can cause failures. Furthermore, clinical use still requires an expert human quality control step for applications such as Neurosurgical planning [56] until the methods have been sufficiently developed and validated. The typical way to select streamlines as part of the bundle-of-interest is to use a tractography output viewer, such as Trackvis [163], to place regions-of-interest (ROIs) manually that select included or excluded streamlines. There are many reproducibility concerns [162, 60] with these methods, however. We propose a cluster-based approach as an alternative to manual placement of ROIs to isolate fascicle models from tractography output. This approach minimizes the variability in manual execution of streamline selection by reducing the output to discrete clusters that require limited decisions for inclusion instead of relying on the placement of ROIs in continuous space. This method also provides the framework for training a classifier that could be tailored to

the data type and goals of a particular application. This is an important consideration, as the tractography output can vary widely depending on a variety of parameters (stopping condition, maximum turning angle, etc.) [39] and there may not be a consensus on what sub-bundles should be included in tractography models for a given application.

## 3.2 Description

This viewer 3.1 enables the user to select streamlines on a cluster-level. The Quickbundles algorithm [67], implemented in dipy (http://nipy.org/dipy/) [66], can be used to quickly cluster a set of streamlines into sub-bundles. The main design requirement for this interactive tool was to minimize the computing time spent reclustering between iterative steps of cluster selection. Quickbundles does not take the computational time needed to optimally cluster streamlines, but rather prioritizes speed to reduce the dimensionality of the classification problem [67]. The user can select all of the sub-bundles that include parts of the target bundle-of-interest 3.2. The user can alternate between selected and deselected streamline bundles by clicking on the button "Toggle Choice" to study the rejected streamlines more closely. The selected sub-bundles are re-clustered into finer sub-bundles when the user pushes the "Finer" button, and the desired components of the bundle-of-interest can be further refined by selecting a subset of the reclustered bundles 3.3.

## 3.3 Results

This Cluster-Based Streamline Tool (https://github.com/kesshijordan/Cluster-viz) was implemented as a web-based viewer with a python backend using CherryPy (http://cherrypy.org/) 3.4. The code from the AFQ-Browser (https://github.com/yeatmanlab/AFQ-Browser) was used as a interface skeleton and adapted for this project. The user can upload and download tractography streamline data, select streamline bundles, and initiate finer clustering using Quickbundles [67]. The viewer presents all of the streamlines to the user and allows them to select a subset of the ten sub-bundles by either clicking on the streamlines, themselves, or by clicking on the menu. This tool is a work-in-progress; in the future, the selected sub-

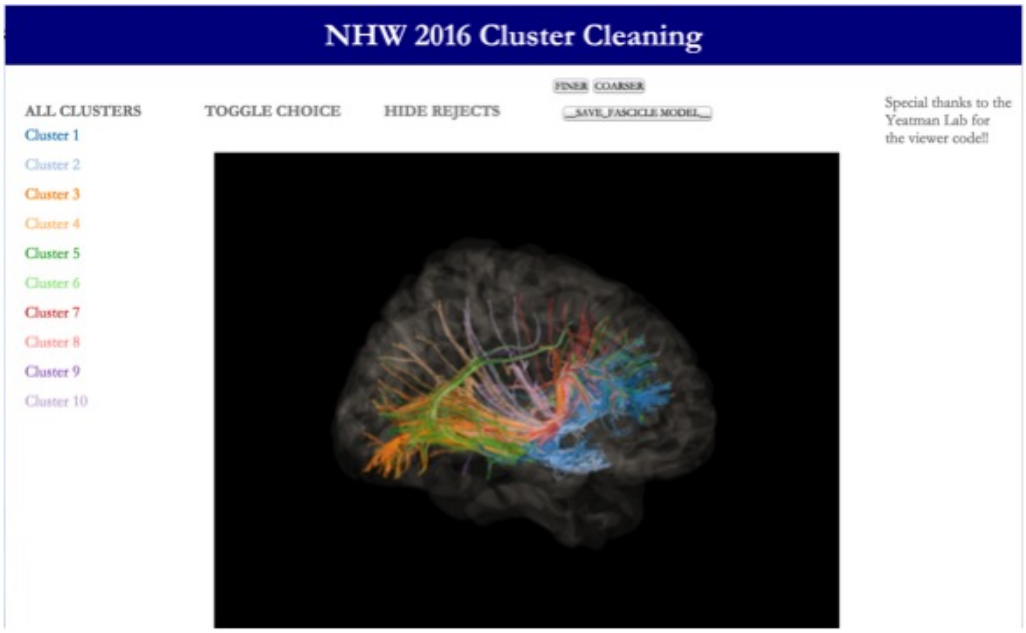


Figure 3.1: Cluster-viz Demonstration

*The connectivity of an ROI placed on the coronal plane over the external/extreme capsules at the level of the anterior commissure is shown (tractography method: Caverzasi et al. 2015). Each color is a cluster, as generated by the Quickbundles algorithm [67]*

bundles will be clustered further upon user request. The transparent cortical surface is for orientation only; it is not in the patient space. In-progress developments include patient-specific anatomical reference in both slice and surface representation and iterative clustering functionality.

## 3.4 Conclusions and Future Directions

This method is advantageous to the traditional ROI-based approach because binary decisions made on discrete clusters is less variable than manually placing ROIs in continuous space. In theory, this should facilitate reproducibility of human operators, as well as create a more tractable training set for machine learning applications. Ideally, the Cluster-viz tool would learn from the user as they interact with the viewer and provide suggestions for bundle classification that the user could approve. Over time, the learning element could greatly

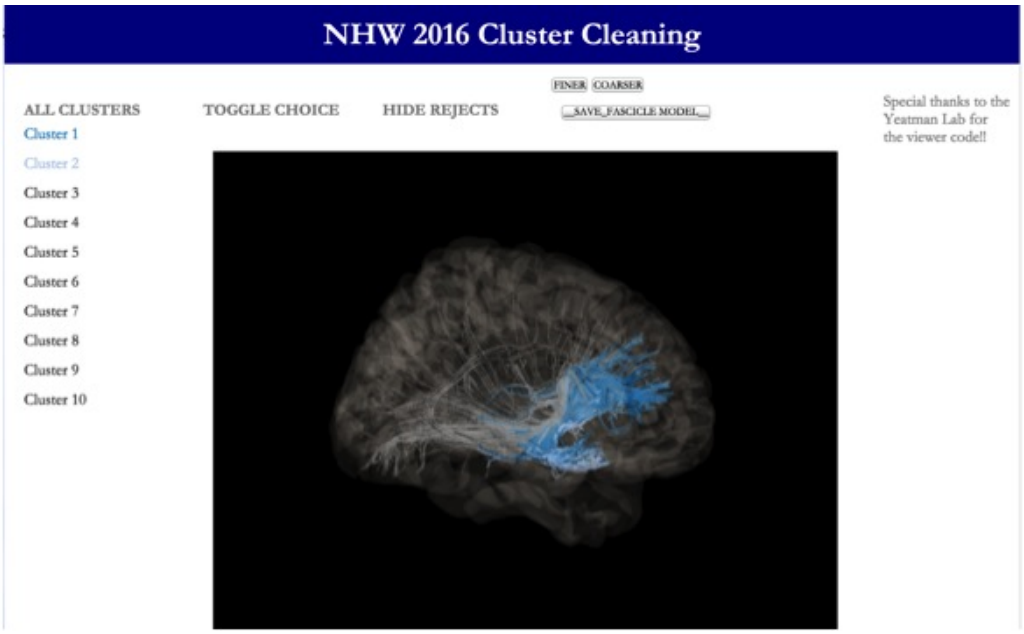


Figure 3.2: Cluster-viz Demonstration Step 1: Select Bundles

*The user selected two sub-bundles that contain streamlines representing a tractography model of the Uncinate Fasciculus.*

increase the efficiency of the user and, perhaps, eventually replace the human.

## 3.5 Author Contributions

This work was conducted in collaboration with Anisha Keshavan, MariaLuisa Mandelli, and Roland Henry.

## 3.6 Acknowledgements

This work was completed during Neurohackweek 2016 in Seattle, WA and the BrainHack 2016 in Los Angeles, CA. We would like to thank Dr. Ariel Rokem and Dr. Jason Yeatman for their help during Neurohackweek and Dr. Jeremy Maitin-Shepard for his help during BrainHack LA. We would also like to thank all of the Neurohackweek and BrainHack or-

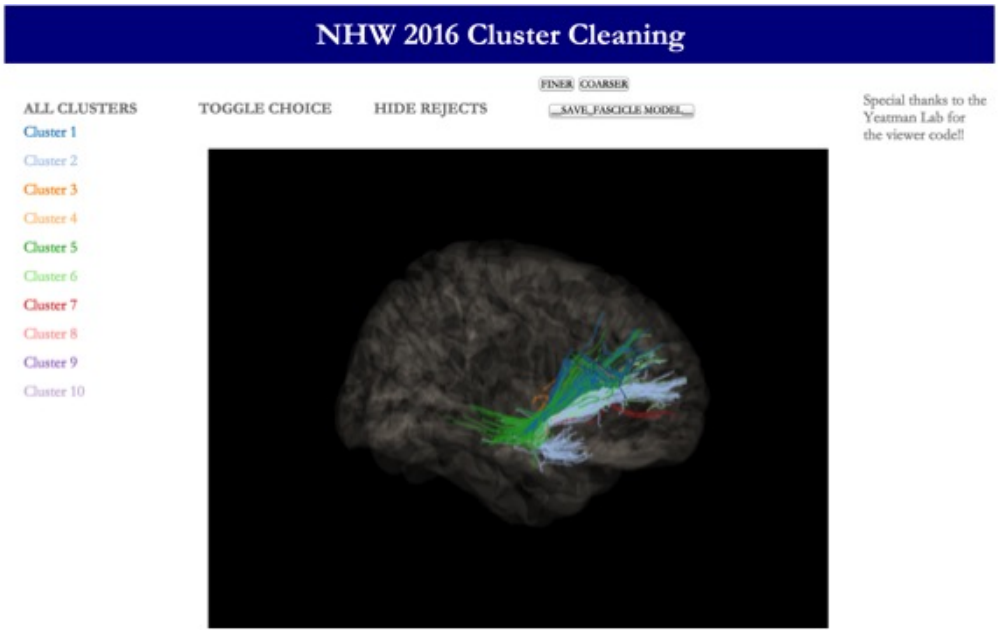


Figure 3.3: Cluster-viz Demonstration Step 2: Recluster

*Sub-bundles that the user judged were part of an Uncinate Fasciculus tractography model are re-clustered so that the user can further refine the model.*

ganizers and mentors. This work was published as a project report in Research Ideas and Outcomes (doi: 10.3897/rio.3.e12394).

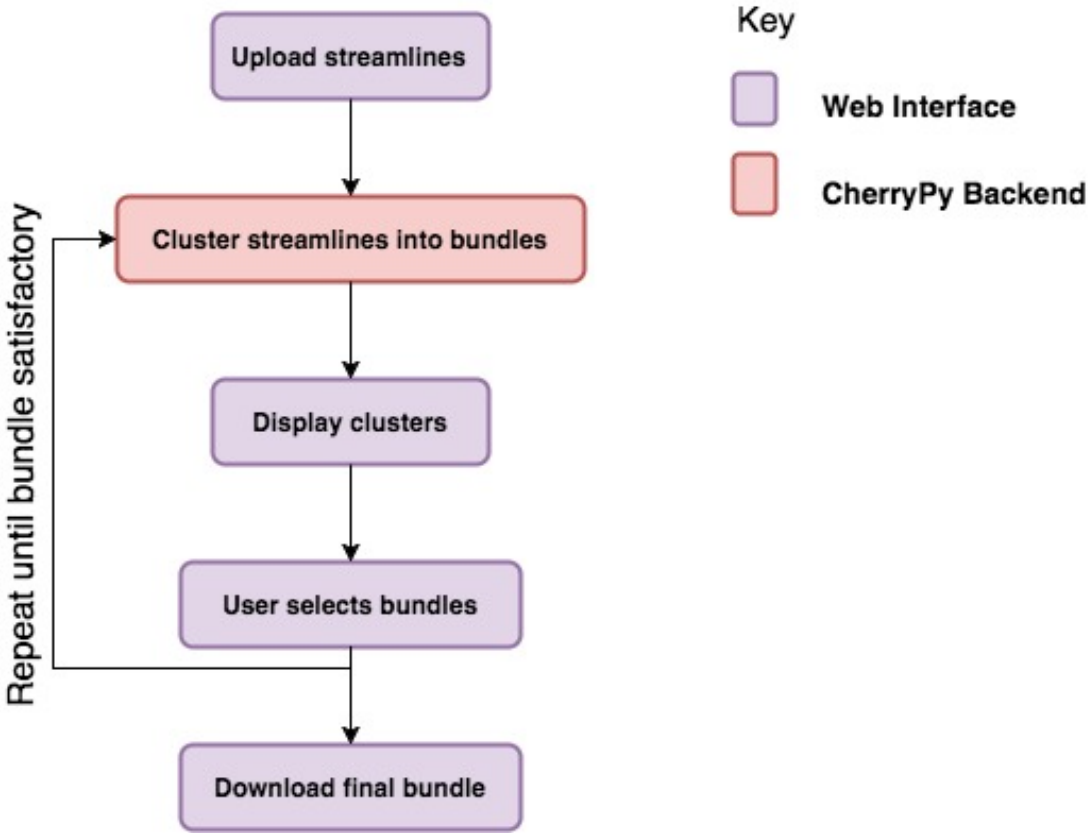


Figure 3.4: Flowchart of the Cluster-viz web application.

# Chapter 4

# An Automatic Pipeline to Create Longitudinal Disconnection Tractograms

## 4.1 Introduction

### Clinical Translations

Elucidating the functional impact of new focal or progressive injury to the brain's intricate communication network is relevant to a wide range of clinical disorders including focal lesions in brain tumors, multiple sclerosis, epilepsy, and stroke, as well as more widespread injuries associated with Tramautic Brain Injury and neurodegenerative disorders (for example Parkinson's Disease and Alzheimer's Disease). This relationship has typically been investigated by either registering an estimate of the injurious region of interest to a template of putative fiber bundles (volume-based method), or by modeling those fiber bundles in each patient using tractography (connectivity-based method). Tractography reconstruction of white matter structures for pre-surgical planning and guidance of intra-operative cortical and subcortical stimulation mapping has emerged as a major clinical translation of diffusion MRI applications. Tractography methods have shown great potential to improve clinical outcome in neurosurgery when used in tandem with Intra-Operative Stimulation

(IOS) [167, 109, 82, 112, 111, 21, 29, 119, 37, 20], but translation to the clinic requires extensive validation [103, 56]. Tractography is an inexact science that is highly dependent on implementation choices [39, 64, 45, 126]. Algorithms with sufficient complexity to model crossing regions also produce many false-positive connections that must be quality-controlled by a human operator, exacerbating well-documented reproducibility concerns [162, 20]. The added complexity of pathology introduces additional errors, which can vary with tractography algorithm choices and underlying anatomy [75, 39]. Establishing robust relationships between deficits and damage patterns using reproducible, operator-independent methods is essential to being able to guide neurosurgical intervention at the single-subject level.

## Volume-Based Approaches

There are several ways to evaluate the impact of *focal white matter damage* on fiber pathways that reflect different levels of efficiency and accuracy. The challenges are (i) segmentation of the focal injury and (ii) association of this focal injury with fiber pathways. The simplest approach is to localize damage on an anatomical image and infer what structures were likely to have been damaged either using a white matter atlas or judgment by an expert (ex: neuroradiologist). A widely used example of this approach is called Voxel-Based Lesion-Symptom Mapping (VLSM), which has been used to great effect in leveraging modern imaging to conduct these lesion studies [12, 104, 31, 3].

Another volume-based method demonstrated by [89] approached the question of relating functional deficit to tissue by examining the residual tumor spared during surgery that was associated with transient functional deficits during direct electrical stimulation (DES). These investigators created an atlas of "functional resectability" by overlaying post-surgical residual tumor from a large cohort of WHO grade II glioma patients to indicate regions that were more or less "resectable" using the gold standard of DES for identifying functional boundaries. These authors chose to acquire post-operative images 3 months post-surgery to minimize the post-surgical displacement of anatomical structures.

While volume-based methods can be an efficient and straightforward approach using conventional imaging, the brains subjected to pathological or neurosurgical damage often will present unique distortions resulting in misregistration to the template. Current methods for

addressing some of these concerns are to exclude the pathology during normalization using a method called cost-function masking [28]. Excluding the pathology from alignment calculation enables registration of distorted brains to a template, but does not attempt to align the distorted region, which may be relevant to the study. Unified segmentation (from the SPM toolbox) has been shown to perform well in stroke patients [7], especially in combination with cost-function masking [5]. However, these methods are highly dependent on a trained neuroradiologist creating a precise lesion mask, which introduces issues in both subjectivity and time efficiency of the procedure. According to [5], this procedure takes about 8 hours per stroke case for a "moderately large lesion", but the combination of unified segmentation and cost-function masking allowed a rougher lesion mask drawing (25-30 seconds per slice instead of 6 minutes) [5]. This approach reduces the time investment considerably [5], but is still a prohibitively labor-intensive endeavor for studies with sample sizes in the hundreds or thousands and does not address the subjectivity/reproducibility issues inherent to manual tracing of focal damage [63] or the need to exclude patients with excessive anatomical distortion (abnormal ventricular size, pneumocephalus, hemorrhage, tumor-associated deformation, etc.). Many alternative methods are being developed to address these challenges [165], but many of these approaches are designed for stroke applications and may not allow wider applicability for cases with increased heterogeneity and complexity, like brain tumors. Tumors are more difficult to develop general methods for because they are infiltrative, meaning that a distinct boundary does not exist at the periphery of the tumor, so abnormal tumor is present within normal brain parenchyma. These inherent properties have varying amounts of heterogeneity, density, and rate of growth, which will alter the mechanics of deformation and appearance on imaging. The task of tumor segmentation is widely acknowledged as a difficult problem, both for consistent manual segmentation and automatic methods [76, 123, 141].

An additional challenge in relating damage to white matter disconnection, in particular, is the lack of white matter contrast in anatomical scans; without diffusion MRI (dMRI), damage to a particular white matter structure cannot be directly assessed. An expert user can infer displacement from anatomical images by estimating mechanical deformation based on prior knowledge and observation of tissue deformation, but the white matter variability

both between people and pathologies makes direct evaluation of the white matter structure advantageous. Tissue displacement can be mistaken for destruction and the highly subjective nature of estimating tissue deformation adds uncertainty to the results. Some approaches perform volume-based methods on anatomical images and extrapolate results to white matter structure by applying fiber tracking in a healthy control cohort to probe white matter connectivity of the resulting regions-of-interest [124] or by overlaying results on a white matter atlas [89], but they are limited by the absence of white matter contrast on the single-subject level. These constraints inherent to any method blind to white matter structure risk biasing the study by misrepresenting or excluding the most extreme cases.

## Connectivity-Based Approaches

A second approach to evaluate focal white matter damage is to reconstruct fascicles using tractography and inferring damage to the structure using properties of the segmented volume (e.g. stroke), or apparent displacement/destruction of the reconstructed model (e.g. tumor). In the neurosurgical example, these studies can be conducted longitudinally because the state of the brain's white matter prior to the injury can be acquired. One way to execute this type of study is to model the fascicle pre-operatively and infer damage using a segmentation of the post-operative resection cavity. Without post-operative white matter contrast, however, it is difficult to determine if bundles-of-interest moved during surgery so adjacent structures could be confounded, and it is impossible to conclusively evaluate white matter registration accuracy. An alternative approach is to reconstruct the fascicle model using tractography at two time points (pre- and post-surgery) and subjectively judge how the integrity of the fascicle changed based on the track model [37]. This method is, arguably, the most direct way to estimate fascicle damage using anatomical imaging, but it is highly subjective both at the step of interactively targeting a collection of streamlines to form the track model and at the rating of track model integrity. This method also requires a priori selection of fascicles that could be involved, since each evaluated fascicle must be hand-segmented into a track model. Hand-segmentation makes this method extremely time-consuming and introduces significant reproducibility concerns, but is necessary because automatic algorithms are not yet robust to many of the deformations characteristic of neurosurgical applications.

## Technical Challenges to Investigating Tumor Resection

Executing group imaging studies in the field of brain tumor surgery is particularly difficult [27]. The heterogeneity of brain tumor pathology, often presenting with solid, infiltrative components and perilesional vasogenic edema, can change the diffusion properties in tissues-of-interest. Moreover, tumor growth can displace or disrupt pathways [62]. Tissue can shift considerably during surgery [127], so spatial relationships between pre-, intra-, and post-operative imaging may not be maintained. Imaging post-operatively to evaluate the effect of surgical intervention has unique challenges, as well. The presence of intracranial blood products, air, and metallic implants used on the surface of the skull in craniotomies can cause imaging artifacts in some modalities. Taken together, these complications cause many of the automatic computational tools used to extract salient features from large neuroimaging datasets to fail, requiring significant human time investment and/or exclusion of cases. As a result, the typical association study between tractography models and deficits is underpowered and the results are often operator-dependent. There is a need for an automatic method to objectively quantify focal white matter injury.

## Proposed Method: Automatic Pipeline to Create Disconnection Tractograms

In this manuscript, we present an automatic pipeline for comparing diffusion MRI volumes and identifying spatially coherent focal white matter disconnection. This pipeline was developed based on a tumor cohort undergoing surgical resection, but the method could be extended to study longitudinal focal white matter changes in many other contexts.

The proposed method (ASAP-tractograms) has two essential components to address the challenges outlined, above. The first component is an automatic segmentation of anisotropic power changes (ASAP), which identifies tissue likely to be associated with damage to long-range pathways based on diffusion properties. This approach completely bypasses the challenges inherent to automatic segmentation and classification of various lesion components to solely focus on focal white matter damage. The second component involves modeling disconnection using tractography on the single-subject level with HARDI-based tractograms.

This approach enables a comparison between subjects that is based on individual characterization of damage to long-range pathways, so comparisons between patients are made based directly on white matter structures. The advantages of this approach stem from (1) objective and automatic lesion segmentation and tractogram generation, (2) objective and precise segmentation of affected tissue likely to be associated with damage to long-range white matter pathways (defined by anisotropic power), (3) good performance even in the cases of anatomical distortions by use of diffusion MRI based registration in the patient space, which aligns images using white matter contrast.

We employ tensor-based registration [173] to ensure alignment of white-matter structures between the pre- and post-operative images. We subtract the Anisotropic Power Maps [49] of the aligned images to determine the regions of steep anisotropy drop, and boost the value of spatially coherent areas [148, 166, 150, 91] to emphasize large regions of pre-surgical anisotropic tissue that were not present in the post-surgical image. We can then delineate a region-of-interest (ROI) that, presumably, represents resected pathological white matter (ASAP ROI). We use residual bootstrap q-ball tractography [23] to infer the disconnection that occurred during surgery by modeling the connectivity of this ASAP ROI pre-surgery.

To evaluate this approach against an established method, we apply this pipeline to a cohort of 35 tumor patients that underwent surgical resection to identify white matter fascicles damaged by surgery and compare the results to a subjective analysis performed by neuroradiologists on the same cohort [37]. This study suggested that preservation of the Arcuate and temporal-parietal components of the Superior Longitudinal Fasciculus/Arcuate complex may be associated with positive outcome, but the study was not sufficiently powered to draw high-confidence conclusions [37]. High-throughput approaches, like the one presented in this paper, facilitate large studies that will empower clinicians to ask the question What is the possible morbidity to this unique patient given the location, size, and extent of this infiltrative tumor? so that they can better counsel their patients and plan interventions to minimize complications and maximize the extent of resection, which will impact patient quality and quantity of life. However, we can only truly answer this question with sufficiently powered studies. The extent of variability due to genetics, learning over a lifetime, varying functional states, pathological processes, the brain's response to pathology [53], and uncertainty in

models of structural/functional systems necessitates studies with considerably larger sample sizes and more specific characterization of disrupted white matter structures than are the standard with current technologies. Our work lays the groundwork for high-resolution, larger scale, sufficiently powered analyses by meeting the need for an automatic approach to objectively quantify white matter disconnection.

## 4.2 Methods

### Subjects

The new pipeline was evaluated on pre and post-surgical resection MRI data from thirty five subjects with glioma brain tumors. Research was performed in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by our institution. The Committee on Human Research at the University of California, San Francisco, approved the study protocol. Written informed consent was obtained from all study participants. Each subject underwent a standard pre-surgical Magnetic Resonance Imaging (MRI) scanning protocol including a High Angular Resolution Diffusion Imaging sequence [23]. This cohort is the same used in a previous publication; details can be found in [37]. Of the 35 patients in this cohort, 6 were excluded for acquisition-related reasons (3 scan-rescan dates greater than a week apart, 2 sequence mismatches, 1 had a major artifact obscuring a large portion of the frontal lobe), leaving 29 patients to be analyzed.

### Diffusion Preprocessing

Diffusion preprocessing was applied to both pre- and post-surgery HARDI datasets in parallel, as shown in Figure 4.1. The datasets were first corrected for motion and eddy current distortion using the FMRIB Software Library (FSL) [91] and the gradient table rotated, accordingly [110]. A tensor model was fit to the corrected HARDI data using the open-source package Diffusion Imaging in Python (Dipy) [66] and the resulting parameters used to calculate fractional anisotropy (FA), mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD). An Anisotropic Power Map was calculated (APM) [49] using the Q-ball

model [50] implemented in Dipy. The resulting map is similar to FA, but the measure of anisotropy comes from the Q-ball model instead of the tensor model, so it should represent white matter better, especially in crossing regions [49]. The B0 image was isolated from the HARDI series and used to create a brain mask by skull stripping the image with FSL's Brain Extraction Tool (BET) [149]. This brain mask was applied to all images.

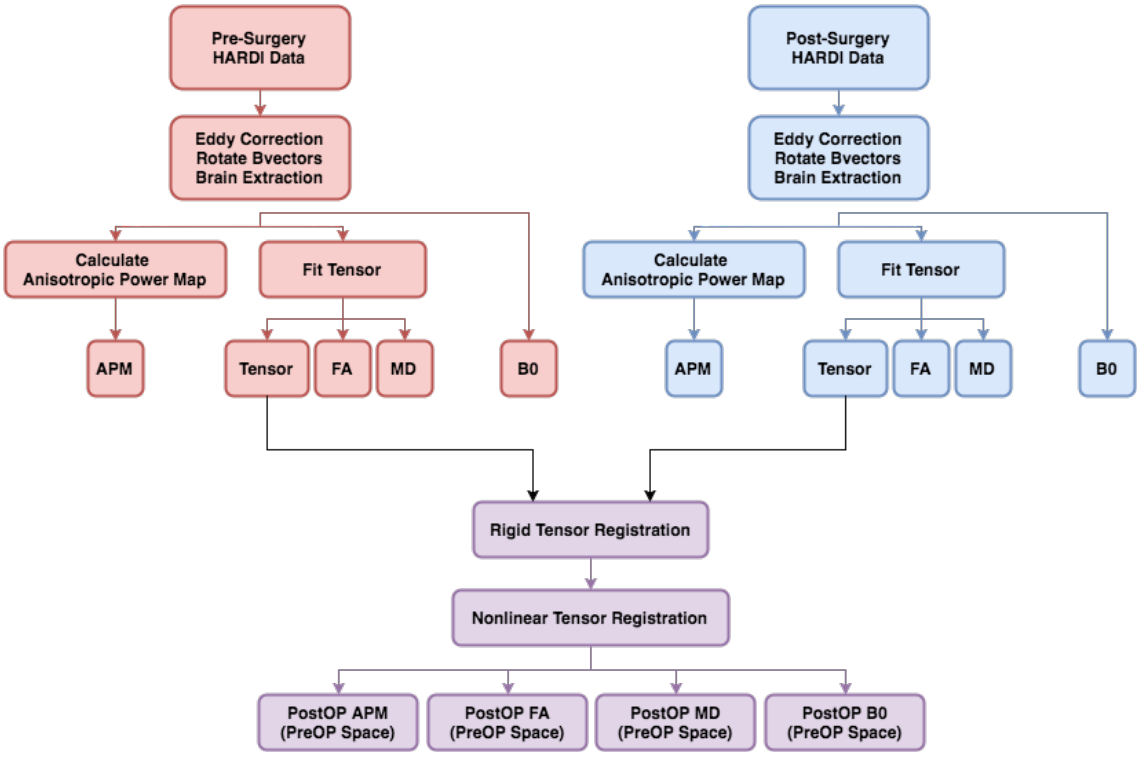


Figure 4.1: Diffusion Preprocessing Flowchart

*A flowchart of the preprocessing steps executed to produce post-surgery diffusion metric images aligned to the pre-surgery images. (APM = Anisotropic Power Map, FA = Fractional Anisotropy, MD = Mean Diffusivity, B0 = T2 image with b-value of zero).*

## Tensor-Based Timepoint Alignment

Rigid and non-linear tensor registration was performed to align white matter structures at the two timepoints using the Diffusion Tensor Imaging ToolKit (DTI-TK) [173]. The transformations were used to bring the post-surgical dMRI images (B0, FA, MD, RD, AD,

APM) into alignment with their pre-surgical counterparts. These images were subtracted pre-surgery minus post-surgery to produce perisurgical difference maps, shown in Figure 4.2. To segment the resection cavity, the flowchart shown in Figure 4.3 was applied. The registered pre- and post-surgical Anisotropic Power Maps were subtracted and spatially clustered using FSL's Threshold-Free-Cluster-Enhancement (TFCE) [148, 166, 150, 91] to boost the signal of spatially coherent drops in anisotropic power. This made it possible to isolate the anisotropic resection ROI using simple thresholds. These TFCE maps were thresholded at the 99th percentile, which was sufficient to isolate the anisotropic resection ROI in most patients, but additional noise from tissue shifting necessitated the additional requirements of a hard threshold TFCE-boosted intensity >500 and a cluster size of >5. This result was quality-controlled using the Mindcontrol platform [101] for accurate representation of post-surgical cavity.

## Tractography

Any tractography method can be applied to produce a prediction of the white matter structures disrupted by the surgery. For the demonstration detailed by this manuscript, the pre-surgical whole brain white matter was seeded (anisotropic power >4) and tracked using residual bootstrap probabilistic q-ball tractography [23] with the parameters described in [37]. Briefly, the preprocessed diffusion signal was fit using spherical harmonics (even orders up to 4) so that orientation distribution functions (ODFs) with constant solid angle factor could be estimated from the data [157, 155, 156]. To perform residual-bootstrap tractography [23], principal fiber orientation was estimated at each step by computing a bootstrapped ODF and identifying the peaks (exclusion criteria for peaks: <45 from a larger peak or peak value <25% of ODF maximum). The principal fiber orientations from the bootstrap ODFs provided the distribution of fiber tracking directions. Tracking was terminated upon the following criteria: FA threshold of 0.15 and maximum angle of 60 [29].

The whole-brain streamline dataset was targeted using the anisotropic resection ROI to infer the connectivity of the resected anisotropic tissue. The resulting distribution of streamlines represents likely pathways of major underlying white matter structures at risk of being damaged by the surgical resection. The pathways with higher confidence have many

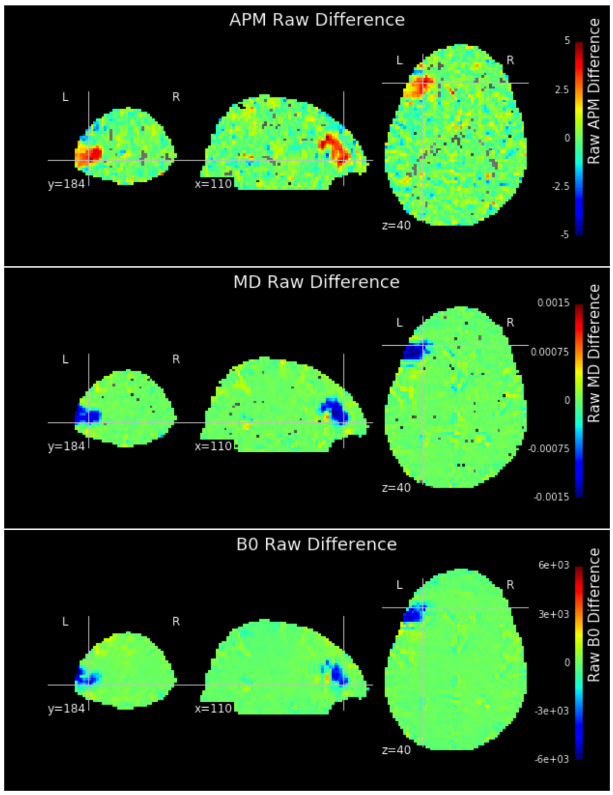


Figure 4.2: Diffusion Metric Difference Map Example

*Difference maps were created by subtracting the aligned post-surgical diffusion metric image from the pre-surgical. The resection cavity is shown as positive in the anisotropic power map (APM) because anisotropic tissue was resected, and negative in the mean diffusivity (MD) and B0 maps because tissue was replaced with CSF, which has higher diffusivity and T2-contrast intensity.*

streamlines following roughly the same trajectory. This confidence can be estimated using the Cluster Confidence Index (CCI) [99]. To eliminate low-confidence pathways that make results noisy and difficult-to-interpret, any streamline shorter than 40mm or with CCI <1 (calculated using default CCI parameters: k=1, theta=5, subsamp=8) was excluded from figures and analysis. This procedure is depicted in Figure 4.4; the code to calculate the CCI is open-source and can be downloaded from GitHub (https://github.com/kesshijordan).

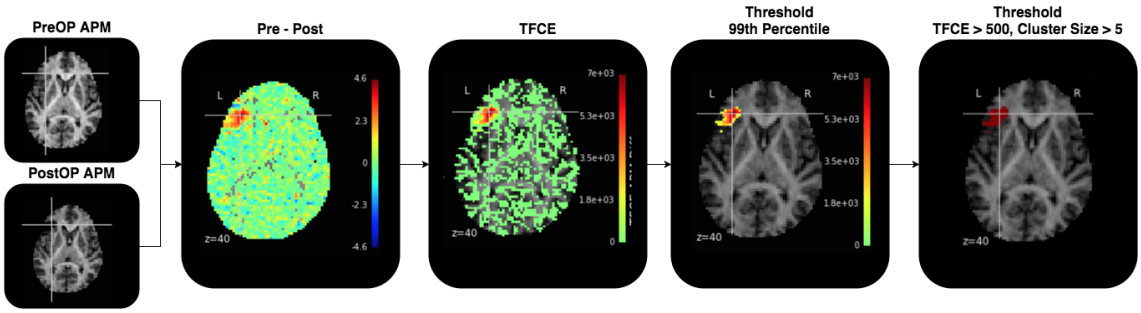


Figure 4.3: Focal White Matter Damage Segmentation Flowchart

*The pre- and the post-surgical Anisotropic Power Maps are subtracted to produce the Pre - Post image. This image is spatially clustered using FSL's threshold-free cluster enhancement to boost the value of spatially coherent decreases in anisotropic power, resulting in the Threshold-Free Cluster Enhanced (TFCE)[148] image. Several thresholds were applied ($99^{th}$ percentile, TFCE-boosted intensity >500, cluster size >5) to isolate the anisotropic resection ROI.*

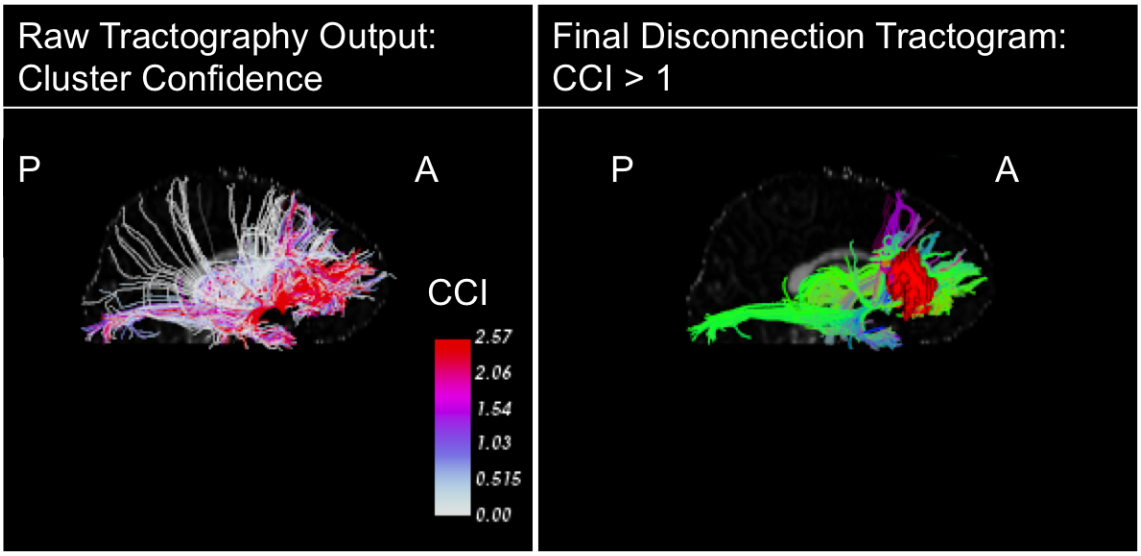


Figure 4.4: Cluster Confidence Index (CCI) Filtering of Streamlines

*The streamline output was filtered objectively using the Cluster Confidence Index (CCI), to exclude low-confidence streamlines (left panel). Any streamlines with CCI <1 or length <40mm were excluded from analysis, resulting in the final disconnection tractogram (right panel), shown with the anisotropic resection ROI.*

### Software Implementation

This pipeline was built and run using the Nipype software package [77]. The resection ROI output was quality-controlled using the Mindcontrol software package [101]. The code is open-source and can be downloaded from GitHub (https://github.com/kesshijordan).

### Comparison to Manual Analysis

The results of this pipeline were compared to those generated by [37]. In [37], the following white matter bundles were segmented independently by two neuroradiologists: arcuate fasciculus (AF), components 2 and 3 of the SLF (SLF II and SLF III), temporo-parietal component of the SLF (SLF-tp), inferior fronto-occipital fasciculus (IFOF), uncinate fasciculus (UF), inferior longitudinal fasciculus (ILF), middle longitudinal fasciculus (Md-LF), according to methods outlined in their cited paper [37, 34, 38, 117], along with the corticospinal tract (CST) and optic radiation (OR) [29] [87]. The same two neuroradiologists independently rated the degree to which each of these fascicles had been disrupted both pre- and post-surgery as unchanged (0), displaced but otherwise normal appearing (1), partially interrupted (2), or completely interrupted (3). These categories were reduced to affected (0 or 1) and not affected (2 or 3). A pre- to post-surgical change from 1 to 2 or from 2 to 3 should be reflected in the results of the difference pipeline because either of those transformations indicate removal of tissue that caused a disconnection.

## 4.3 Results

### Evaluation of Performance: Connectivity of Anisotropic Resection ROI

The streamline output of the difference pipeline clearly shows the white matter structures hypothesized to have been disrupted by the surgery (Figure 4.5). These sub-bundles are recognizable by shape and their pathway through the brain. For the purposes of demonstrating this using a two-dimensional figure, the streamline output for each case has been colored by a researcher with experience in fascicle modeling to highlight the bundles associated with

motor, optic, or language that are routinely modeled for pre-surgical planning research at UCSF (Figure 4.6). These illustrative cases are compared to the subjective rating conducted by [37]. This study manually segmented fascicle models of bundles pre- and post-surgery, rating each subjectively on a scale from 0 to 3 (0=unaffected, 1=infiltrated/displaced, 2=partially destroyed, 3=completely destroyed). The difference pipeline shows disconnections in tractography models, so we would expect bundles output by the difference pipeline to reflect rating changes from 0/1→2 or 0/1/2→3. In theory, the rating scale could be unchanged between the pre- and post- surgery in the presence of resection-related disconnection because a fascicle that was partially destroyed pre-surgery and further damaged by the resection would receive an identical partial destruction rating (score of 2) for both time points. This rating scale is not resolute enough to specify extent of damage beyond partial to full disconnection. Changes 0→1 or 1→2 would not be reflected in the difference pipeline results because they do not indicate disconnections.

Figure 4.5 demonstrates the basic idea: the left panel shows the connectivity of the difference pipeline output anisotropic resection ASAP ROI (colored red), as modeled by tractography. We can see a bundle that follows the path of the Uncinate Fasciculus (purple) from the frontal lobe, hooking down and around into the anterior temporal lobe and another bundle that follows the path of the Inferior Fronto-Occipital Fasciculus (green), from the frontal lobe through the external/extreme capsule and continuing posteriorly to terminate in the occipital lobe. The connectivity of the anisotropic resection ROI suggests that these two bundles were disconnected by the surgery; specifically, the sub-bundles of the UF that projected laterally into the frontal cortex and possibly some UF components that terminate anteriorly in the frontal lobe. The right panel shows the pre- and post-surgery tractography models from which the damage rating was assigned. From these manual reconstructions, it appears that the lateral projections of both the UF and IFOF present in the pre-surgery model are missing from the post-surgery model. In this case, the subjective damage rating of the UF and IFOF increased from 1 (infiltrated/displaced) to 2 (partially destroyed), so the approaches reached the same conclusion with respect to these two structures. The subjective rating also scored the SLF II as increasing from a score of 2 to 3, but the resection did not overlap with the missing streamlines from the SLF II (the underlying white matter appears

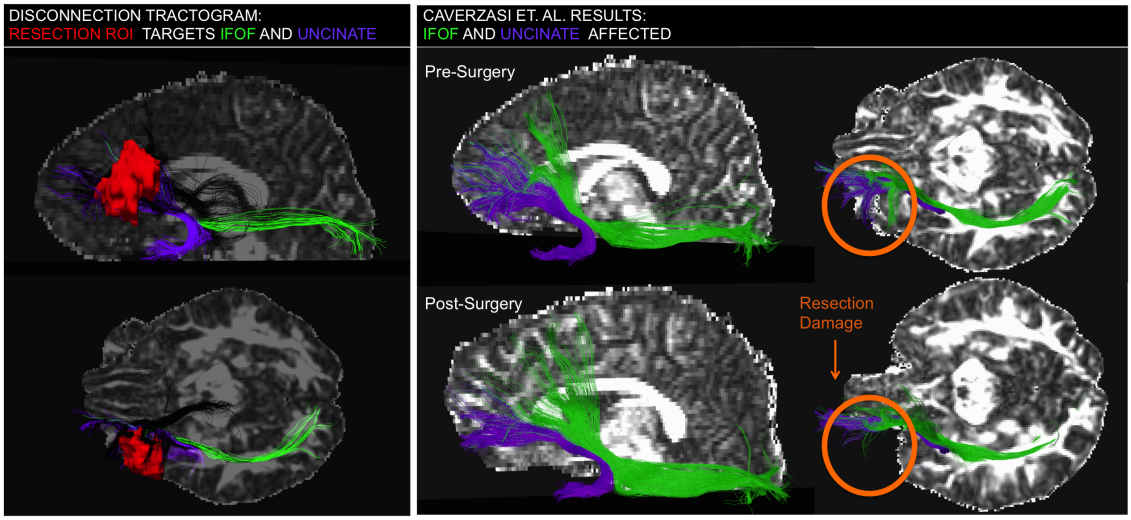


Figure 4.5: Disconnection Tractograpm vs. Pre-Post Surgery Analysis

*Comparing Difference Pipeline to Pre-/Post-Surgery Anatomically Constrained Fascicle Models LEFT: The Difference Pipeline Results suggest that the resection ROI (red) disconnected segments of the Inferior Fronto-Occipital (green) and the Uncinate (purple) Fasciculi. The RIGHT panel shows the manual/subjective comparison (Caverzasi 2016) study results pre (LEFT) and post (RIGHT) surgery in the transverse plane. The orange circle indicates the change pre- vs. post-surgery. The manual results indicated potential disruption of the SLFII, but the underlying white matter is intact so it is likely due to post-surgical tracking difficulty should not be included in the disconnection tractogram. Manual Results: IFOF (1→2), Uncinate (1→2), SLFII (2→3), SLFIII (2→2), Arcuate (1→1)*

intact). It is possible that this discrepancy is due to tracking difficulty in the post-surgical dataset. The presence of edema has the potential to influence these tractography models significantly because the stopping condition is based on fractional anisotropy, which drops in the presence of edema. This case demonstrates an important advantage of the difference-pipeline approach: each anatomically-defined fascicle is made up of a variety of sub-bundles that project to different gyri and may support different functions or have different levels of redundancy. The difference-pipeline approach indicates that the sub-bundles of the UF and IFOF projecting laterally are likely to have been disrupted, which could be a relevant distinction to make when investigating the relationship of white matter disconnection to functional deficits in a patient. Also, the differences in tractography conditions between the pre- and post-surgical datasets is less problematic because the tracking is all performed in

the pre-sugical dataset to create the disconnection tractogram.

Disconnection tractograms are shown for the entire cohort that passed quality-control (26/29 subjects) in Figure 4.6. The left image for each patient shows the automatically-generated connectivity of the anisotropic resection ASAP ROI (CCI >1), as described in the methods, colored by standard orientation. In the standard orientation color-scheme, each streamline is colored according to the dominant trajectory, as seen on the traditional FA color map (red=right/left; blue=superior/inferior; green=anterior/posterior). The right image shows any fascicles-of-interest colored according to the provided key. Of the 29 patients that met inclusion criteria, 3 did not pass quality control: one patient had a bad registration between the pre- and post-surgical images and 2 had false-positive ASAP ROI's present. In the 3 cases investigated for anomalous large regions of Anisotropic Power increase post-surgery, other imaging contrasts indicated postoperative blood products and/or pneumocephalus colocalizing with the region.

## Automatic Pipeline

One of the most obvious advantages to an automatic pipeline is labor-cost savings. This cohort was processed with minimal intervention by a human operator. All of the pre-processing, image registration, difference map calculation, clustering, and thresholding was handled entirely by the pipeline code. Depending on how the tractography is implemented, that step could be easily integrated, as well. The human requirements to generate a disconnection tractogram are quality-control of the pipeline inputs and output, which involves quickly looking at a pre-surgical diffusion image, a post-surgical diffusion image, and overlaying the ROI. The streamline connectivity of the anisotropic resection ROI output could, in theory, be generated completely automatically and be sent to a neuroradiologist for quality control (QC) and reading of which white matter structures were at risk of being disrupted by the surgical resection. To demonstrate: if we look at the disconnection tractograms from all of the patients in this cohort that passed QC in Figure 4.6, it is straightforward to identify what bundles were at risk of being damaged by the surgical resection, especially in cases U→Z, which show no streamlines belonging to fascicles-of-interest. Using a manual approach would require all of the surrounding fascicle models to be tracked and segmented to judge whether

or not they had been damaged by the surgery. Using this disconnection tractogram generated from the difference pipeline, we can easily see that no fascicles-of-interest were included in the disconnection tractogram and conclude that none of the fascicles in our study were likely to have been damaged by the resection.

## Continuous Representation of White Matter Damage

One advantage of representing white matter damage continuously, as opposed to using a rating scale, is that some damage patterns cannot be described discretely. Figure 4.7 demonstrates this point. From the connectivity of the anisotropic resection ROI (Figure 4.7 left panel), we can segment portions of the left Arcuate Fascicle (yellow) and left Inferior Fronto-Occipital Fascicle (green). This agrees with the manual reconstructions of the pre- and post-surgery Arcuate and IFOF (Figure 4.7 right panel). The IFOF appears to have lost a large portion of its superior-frontal lobe connections as a result of the resection, which is reflected in the damage rating increase (rating 1→2) and noted in the resected tissue missing from the post-surgery IFOF model (orange arrow). The Arcuate, however, did not change in subjective rating pre- vs. post-surgery (rating 2→2); the fascicle model was partially destroyed pre-surgery and futher destroyed post-surgery, but the rating system was not resolute enough to reflect that distinction. This case shows the advantage of having a continuous representation of damage, as opposed to a discrete one; the Arcuate is subjectively codified as unchanged, but the difference pipeline connectivity shows the specific connection that was threatened by the surgery.

Another case that demonstrates the advantage of evaluating white matter disconnection on a continuous scale is shown in Figures 4.8 and 4.9. According to the disconnection tractogram, the region immediately adjacent to the tumor is connected to the temporal lobe, frontal lobe, and spinal cord. However, these streamlines do not meet the criteria for inclusion in the any Anatomically-Constrained fascicle models because both ends do not reach the cortex. These bundles follow trajectories characteristic of the Arcuate Fascicle (yellow), the IFOF (green), and a bundle connected to the spinal cord (blue). The pre-surgical model of the Arcuate does not include the component wrapped around the tumor because it does not reach the cortex, so the pre- vs. post-surgical models were rated as

unchanged.

## Performance Despite Distortion and Artifacts

The difference pipeline performs well, in spite of interference from artifacts or tissue distortion. As we can see in Figure 4.10, the pipeline output suggests that the Arcuate (yellow), Medial Longitudinal Fascicle (MdLF=orange), the Inferior Longitudinal Fascicle (ILF), and a small bundle going to the spinal cord (blue) were at risk of being damaged by the surgical resection. The artifact appears to be associated with blood products and pneumocephalus (by comparison to T1 and CT), and was not considered in the ASAP ROI because the anisotropic power increased, instead of decreased. These results reasonably agree with the analysis by [37]. The Arcuate was rated as increasing 1→2 and the MdLF increased from 1→3. The cortico-spinal track was rated as unchanged (1→1), but the difference pipeline result was wrapped around the tumor and did not terminate in the motor cortex, so it would not have been included as part of the pre-surgical CST model. The ILF numerical rating was unchanged (2→2), but the rating scale was not sensitive to severity of partial disconnection, so this does not disagree with our pipeline output results.

Another case that demonstrates the power of this approach in studying cohorts with tissue distortion is shown in Figure4.11. This patient had a particularly septated heterogeneous lesion with cystic, nodular and hemorrhagic components, which was further complicated by post-surgical hemorrhage exerting pressure on the tissue. A mass effect is apparent in the pre- and post-surgical images (Figure 4.11), but the pipeline still resulted in a reasonable output that did not disagree with the manual assessment by neuroradiologists. The output showed a sub-bundle of the Arcuate (yellow) as the likely white matter disconnection occurring as a result of the resection. The manual approach rated the only changes as Arcuate Fascicle and SLF II increasing in damage from 1→2. One point to note, in a case like this, performing tractography on the post-surgical dataset may be very difficult because of challenges associated with performing tractography post-surgery (e.g. edema, blood products, metallic implants, etc.). The cavity is not adjacent to the SLF II, so the damage rating to the SLF II structure is questionable; it is possible that the manual reconstruction was missing components due to post-surgery tracking difficulty, as opposed to the tissue being resected.

The difference pipeline produces disconnection tractograms using pre-surgical data; the post-surgical data is only used to identify cavity, so surgery-related edema or blood products has less effect on the results.

### False Positive Resection ROIs

Several patients failed QC due to the presence of false-positive resection ROI's. One example of this is shown in Figure 4.12: this patient had a false-positive error in the automatic segmentation of the anisotropic resection ROI due to excessive distortion of ventricles. Blood products and pneumocephalus appear to compress the brain postoperatively and cause the ventricles to shift, in relation to the preoperative image. The peri-ventricular white matter is being mistaken for a resection ROI because parts of the highly anisotropic corpus callosum overlay the shifted ventricle in the post-operative image. This caused two patients in the cohort to fail quality control.

## 4.4 Discussion

In this manuscript, we present a pipeline that objectively evaluates focal white matter damage to give researchers a tool for robustly and efficiently advancing the translation of tractography-based methods to the clinic. The need for technical development of tractography methods is widely acknowledged [56]. A study by [60] evaluated the anatomical accuracy of fornix tractography across many algorithms based on anatomical accuracy and number of fibers outside of anatomical boundaries. Based on their findings, the authors stressed that the manual intervention commonly employed to "clean up" false positives with placement of exclusion ROI's creates a danger for excluding displaced fibers that represent genuinely displaced white matter bundles [60]. We found this concern to be justified in this study, as well. As an example, the case presented in Figure 4.8 had subjective ratings of fascicle damage that did not explain the permanent new language deficit that the patient experienced. An analysis of the difference pipeline results suggested that a fragment of the model of the Arcuate Fascicle, a white matter structure associated with language deficits [24], wrapped around the tumor and had potentially been disrupted by the surgery. While this wrapping

phenomenon is not well-characterized, it is reasonable to assume that information can be gleaned from structures engaged in wrapping merely because of their proximity to the tumor, which puts them at risk of being damaged by the surgical resection. In the manual fascicle modeling of this case, streamlines that wrapped around the tumor region were removed by the manual operator because they did not match the anatomical constraints of the Arcuate Fascicle, so the potential disruption was not reflected by the post-operative rating scale because the at-risk bundle was excluded from the pre-operative model 4.9. In the case shown in Figure 4.8, a distorted fascicle model of a bundle indicating connectivity to the spinal cord was wrapped around the tumor, but did not reach the cortex so the identity is inconclusive. The patient had a short-term post-surgical motor deficit, so information about connectivity to the spinal cord may be relevant to this case. It may be that the white matter directly adjacent to the tumor region helps to explain the discrepancy between tissue damage and functional deficits, but tractography models in that tissue are highly uncertain due to the presence of disease processes. The field needs to conduct studies with hundreds (or even thousands) of patients using objective, reproducible methods before the technology can be explored as a tool to inform clinical decisions. Developing methods to minimize the extent of manual intervention is a crucial step toward this goal.

## Advantages

The difference pipeline approach, presented here, is advantageous in several respects. The objective and automatic aspects have vast implications on scalability of studies and translation across institutions, but the pipeline approach, itself, has unique advantages for studies of any size. As demonstrated by Figure 4.7, sub-bundle resolution is important in executing specific studies relating white matter connectivity to functional deficits; the classification of a subcortical pathway may include a variety of connections and there may not be a consensus on the exact components of an anatomically-defined fascicle model. The Arcuate, for example, has terminations on entirely different gyri, each of which may serve distinct functions [61]. The connectivity of the Uncinate Fasciculus is still controversial, but there is evidence for terminations all the way down the orbito-frontal cortex, adjacent to the amygdala, and connections that turn anteriorly into the anterior temporal lobe [79, 81]. Also, there is evi-

dence that a bundle's robustness to damage is not uniform [83]. The controversy surrounding sub-bundle functionality evidences the need for tools that facilitate investigation of white matter structures with sub-bundle resolution. In Figure 4.13, the bundles associated with transient (Arcuate and ILF) and permanent (Arcuate, SLF-ip and SLF-tp) language deficits in this cohort are readily apparent.

Another advantage of an automatic pipeline is minimizing the influence of human subjectivity on results, which present challenges in translating results across studies and institutions. In the [37] study, the raters were presented with ambiguous cases in which the part of the fascicle model typically related to language function had been disconnected, but not the entire anatomically-constrained fascicle model. They had to choose whether to rate the track as "completely destroyed" in the spirit of the study's intended purpose of relating track model damage to functional deficits, or "partially destroyed" using strictly anatomical definitions of fascicle models. There is not yet consensus on the functioanl role of each sub-bundle in most fascicle models, or even what sub-bundles should be included in an anatomically-constrained model (not to mention variability in what these models look like, depending on tractography choices such as tensor vs. Q-Ball model). Subtleties such as these speak to both the need for continuous representation of bundles to distinguish between different sub-components, and the importance of eliminating human intervention to enhance reproducibility and make translation of results and methods across institutions more realistic.

This pipeline approach is also very simple, in terms of imaging. The entire procedure was performed using solely HARDI images (T2, AP, MD, L1, RD are all derived from HARDI data), so there are no inter-modal registration issues. Approaches using T1 anatomy to study diffusion imaging run the risk of mismatch between sequences. Also, the approach only attempts to precisely register images longitudinally; there is no registration to MNI, as is the case in voxel-lesion symptom mapping (VLSM). Cases, such as Figure 4.12, may be excluded from VLSM studies due to the extreme level of tissue distortion. In a study by Herbet and colleagues, "patients with abnormailities on MRI (e.g. tumour-related deformation, hyroma, abnormal ventricle size, etc.) were also excluded... to avoid normalization problems" [83]. These abnormalities are not a problem for this difference pipeline approach, provided that they are consistent. The distortion between pre- and post-surgery are similar enough that

the patient was successfully analyzed by the difference pipeline. The resulting tractography model generated on the single-subject level can be compared in a cohort study without ever registering the distorted tissue, itself, to a common space. Conclusions made using VLSM must be qualified by statements such as "a large proportion of the significant voxels identified in the VLSM analysis... seemed to overlap with the spatial location of the IFOF" [3]. This pipeline models the connectivity of the resection tissue in the patient-specific space, which confers confidence in the identify of the white matter structure threatened by the surgery. Additionally, compared to the [37] approach of using post-surgical tractography models to depict white matter damage, this method is less sensitive to immediate post-operative challenges to tracking, like edema or blood, because all of the tracking is performed in the pre-operative condition. Only regions with the highest, most spatially coherent, drop in Anisotropic Power are included in the analysis.

## Considerations

While this pipeline confers many advantages, several considerations should be taken into account. Some studies have shown that it is important to consider the location of damage along the length of the track. In [83], their results showed varying levels of "neuroplastic potential" in different segments of tracks (for example anterior vs posterior ILF). It is possible that these differences can be attributed to sub-tracks not distinguished by tractography, but it may also be that the characteristics of a track genuinely vary along its length. If the former, then higher-resolution tractography may help address this concern; if the latter, however, it will be important to account for specific location along the length of the track, perhaps using some sort of streamline template-matching.

An element of post-surgical track damage that is not accounted for by this pipeline is ischemic injury. It may be possible to optimize this pipeline for ischemic injury using the difference maps; known changes in MD could be leveraged in a similar manner to how anisotropic power was used in this study by identifying spatially coherent regions of MD decrease between pre- and post-surgical images.

Another consideration when deciding on an experimental approach is when and how to collect the HARDI datasets. There is a tradeoff between some of the challenges present im-

mediately postop, and the risk of tissue reorganization or further disease processes occurring at a later time-point, after immediate post-operative challenges have resolved. Analyzing immediate post-operative images has its challenges, but this approach isolates the focal white matter damage due to surgery. It is possible that serial HARDI scans would be beneficial to address the challenges associated with both time points. There are ways to improve data quality to address some of these post-surgical challenges, however. One of the aspects of post-surgical images that makes registration to pre-surgical images difficult in this particular cohort is the susceptibility artifacts caused by blood, air, and implants being imaged using an Echo-Planar sequence. The HARDI datasets from this study were not collected with any components necessary for susceptibility correction at the time of acquisition. Susceptibility correction can be trivial, provided either a field-map [95] or two B0 images with opposite phase-encoding gradients are collected at the time of acquisition [6]. Applying this pipeline to susceptibility-corrected images should improve performance, as that removes one element of mismatch pre vs. post-surgery. The authors strongly recommend that, if possible, susceptibility correction is included in sequence planning.

## Anisotropic Power Difference Map Artifacts

In several cases, the pattern of anisotropic power decreasing and mean diffusivity increasing pre- to post- surgery was reversed in some regions around the surgical site. Anisotropic power is extremely hyperintense in these regions post-surgery (greater than the intensity of the corpus-callosum, which should have some of the highest anisotropic power in the whole brain)(Example: Figure 4.11). Based on comparison by a neuroradiologist to CT images, there were two explanations for this phenomenon: pneumocephalus and blood products. When air is present inside the skull (pneumocephalus), MD decreases because CSF/tissue is displaced by air, so the anisotropic power presents as artifactually hyperintense post-surgery resulting in a net increase in the AP difference parameter. This parameter reversal was apparent in several cases, as can be seen in Figure 4.10. AP is increasing and MD is decreasing, because anisotropic power appears to be artifactually hyperintense in the presence of blood. The decrease in mean diffusivity (MD) is logical because blood is very viscous, restricting diffusion. Anisotropic power (AP) is a relatively new measurement, so it

has yet to be characterized. This artifact in AP hyperintensity may be due to the dephasing of signal in the B0 image due to inhomogeneity of fields characteristic of (blood because of its high paramagnetic properties and air/tissue interface creates a susceptibility artifact). The diffusion weighted signal (S=Sdwi/Sb0) is equal to the diffusion weighted image divided by the B0 image, so if the B0 image has signal dropout, noise in the diffusion-weighted image is divided by zero and results in an artifactually high AP.

## Tractography Limitations

In one case, the difference pipeline missed some bundles because they were not represented in the pre-surgery streamline dataset used for targeting the resection ROI. This is due to insufficient seeding density; these streamlines were represented in the [37] because the seeding density was much higher. The seeding scheme could be inflated using a method like the iterative Cluster Confidence Index approach [99] or a method like Linear Fascicle Evaluation [132] could be used to identify cases in which the tractography model does not represent the underlying diffusion signal adequately in the pre-surgical tracking. In cases with a high level of edema, the stopping condition should be changed from FA $<0.15$ to an approach more favorable to tracking through edematous tissue would help to ensure that pre-operative bundles are more faithfully represented.

## Challenges Analyzing Resections

Resection studies are very difficult, from an engineering standpoint, because the mechanics of tissues are difficult to predict [52, 158], especially when a heterogeneous pathology is involved [65]. There is a significant concern for how well the tissues can be matched from pre- vs. post-surgery imaging and at what point correcting for the tissue shifts that come with focal lesion resections (the previous tumor growth exerting local pressure that subsequently relaxes back into the resection cavity after removal, which eventually disappears), gravity causing tissue to "cave in to larger resections, blood causing both imaging artifacts and, in some cases, exerting pressure on the brain, implants causing imaging artifacts, air causing both tissue displacement and imaging artifacts, pressure/CSF changes, swelling of tissue, to name a few. Setting parameters for longitudinal registration between perisurgical timepoints involves a

tradeoff: the nonlinear registration algorithm must reflect genuine tissue shift associated with the surgery, but it it cannot be allowed so much flexibility that it "corrects" the surgical cavity. For example, in some cases there are ventricular distortions that are of the same magnitude as resection ROI's so they are manifesting as false-positives (Figure 4.12).

Many concerns about using this ASAP pipeline approach can be addressed by rigorous implementation standards: a pilot study should always be performed with the results from a variety of extreme patient conditions evaluated by neuroradiologists, and a strict QC protocol employed. Several features of diffusion parameters lend themselves to serving as indicators for QC. Pneumocephalus and blood products are easy to flag for QC because inverse parameter changes are straightforward to identify automatically. The presence of blood products could cause tissue damage by a mechanism different from that of resected tissue. By flagging inverted parameter changes, blood products can be treated as a separate group or an expert can decide on a case-by-case basis whether or not to include the voxels in the ASAP ROI as damaged white matter based on other factors (e.g. other imaging or knowledge of surgery). One of the weaknesses of the ASAP pipeline applied to a neurosurgical cohort is a tendency to report false-positive ROI's around the ventricles when large shifts in the brain occur. Many QC approaches could be implemented to address these problems: screening for multiple ROI's a minimum distance apart, using other diffusion parameters that identify the cavity location, or even explicitly screening for periventricular resection ROI's using a rough atlas are all viable options. Pneumocepahalus exerts pressure on the brain, and might indicate more severe post-operative tissue distortion (compared to pre-surgery) and will also cause susceptibility artifacts due to the air-tissue interface that may be useful to flag. Several artifacts appear to manifest as anisotropic power hyperintensity, so cases that may not have been processed as expected can be flagged easily using this metric.

## Potential

In this demonstration, the output of the fascicle model was classified by an expert based on the trajectory of the pathway. Automatic classification algorithms are being advanced all the time ; some have even been demonstrated as robust to pathology [169, 128]. Given the uncertainty associated with reconstruction of fascicle models in the face of extreme

anatomy distortion, it is unlikely that automatic algorithms will be able to replace expert input, entirely. However, it is possible that automatic classification of extreme deformation pathologies is on the horizon, at least for research purposes. This would be a great tool in facilitating large-N fully-automatic studies and (after rigorous testing and engineering development to advance reliability and establish fail-safe protections) be a candidate for translation to the clinic as a resource for informing clinical decisions. With studies leveraging resolution of white matter damage on a continuous scale, maps that show the risk of a particular deficit could be developed with the confidence that comes from very large studies.

The next steps in this work are to 1) optimize this difference pipeline to identify ischemic injury automatically, 2) isolate the pre-surgical damage by comparing the pre-surgical timepoint to a healthy template, and 3) leverage streamline classifiers and fascicle templates to automatically quantify what percentage of a fascicle model any given bundle represents. With these advancements, we would have a tool that automatically quantifies the fascicle damage exerted by each stage of pathology/intervention. This would provide a more complete picture of each patient's disease/intervention course and enable studies accounting for more sources of white matter disconnection. For example, a patient that presented with a tumor, underwent a resection, experienced a recurrence, and underwent a second resection could have a complete story of white matter damage if they receive HARDI scans in conjunction with typical check-up scans. The ASAP-tractogram pipeline could be applied between adjacent timepoints to create a damage trajectory that could then be associated with quality-of-life changes.

The HARDI sequence only adds an additional 6-8 minutes (or less with Multiband capability) to the MRI scan time and has the potential to provide a wealth of information on what white matter damage patterns can be recovered from and what constitutes a permanent functional deficit. After large, reproducible studies are conducted that relate damage patterns to deficits with tight confidence intervals, it may be possible to use these tools to help the patient to make decisions about how much risk they are willing to take on to functional systems and use this information to help plan a surgery that the patient is most comfortable with.

This pipeline could also be applied to other types of longitudinal white matter damage.

Neurosurgical intervention in pathologies involving less tissue distortion should work at least as well as this tumor demonstration (the pre-surgical brains in non-tumor epilepsy patients should be more similar to controls than, for example, a grade IV tumor). Pathologies like stroke and Multiple Sclerosis may require tuning of parameters but, in theory, should be a good application for this pipeline because the deformations between timepoints are much smaller than those associated with tumor resection.

## 4.5 Conclusion

A shift in the framework of thinking about the brain has been progressing from a topographical approach to one more focused on pathways ("hodotopical") [32]. This is reflected in efforts by the field of neurosurgery to move beyond preservation of discrete "eloquent regions" toward a holistic approach integrating information from many anatomical/functional systems [54]. Tractography methods are being pushed into the neurosurgical clinic because they provide invaluable in-vivo information about patient-specific white matter structures, essential to understanding this new paradigm. There is growing evidence that the potential of functional compensation for white matter damage is limited (LGG [89] stroke [80] TBI [69] [46]) and that, for post-surgical plasticity to be possible, long-range white matter connections must be spared by surgery [57]. There is evidence that tissue loss can be compensated for by adjacent regions [18] and that long-range connections may help cortical reorganization [131]. All of these questions are essential to answer as the field considers using diffusion imaging to guide treatment, but the studies required to develop confidence in the answers cannot be done with today's technology. To further advance studies relating focal white matter damage to deficits, we need methods that are objective, automatic, and compatible with brains distorted by pathology.

The ASAP-tractogram approach proposed in this manuscript may provide advantages over current methods for both accuracy and efficiency when confronted with pathology that significantly distorts the normal anatomy (as is often the case for brain tumors). Furthermore, automatic and objective features of the pipeline for both lesion segmentation and tractogram generation may offer advantages even for cases accessible to other approaches. The

proposed method has two essential components, (i) automatic segmentation from anisotropic power (ASAP) differences, and (ii) modeling disconnection using tractography on the single-subject level, which specifically identifies the disconnections associated with focal white matter damage. The advantages of this approach stem from (1) objective and automatic lesion segmentation and tractogram generation, (2) objective and precise segmentation of affected tissue likely to be associated with damage to long-range white matter pathways (defined by anisotropic power), (3) good performance even in the cases of anatomical distortions by use of DW-MRI tensor-based registration in the patient space, which aligns images using white matter contrast.

## 4.6 Author Contributions

This work was conducted in collaboration with Anisha Keshavan, Eduardo Caverzasi, Joseph Osorio, Nico Pappinutto, Bagrat Amirbekian, Mitchel Berger, and Roland Henry.

## 4.7 Funding

This work was supported by the National Institutes of Health [5R01NS066654-05]. KJ was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

## 4.8 Acknowledgments

Thank you to the Neuroimaging in Python (NIPY) development community, all of our research subjects, Ariel Rokem, Francesco Sanvito, Simone Sacco, Antonella Castellano, Valentina Panara, Robert Knight, Shawn Hervey-Jumper, and Vanitha Sankaranarayanan

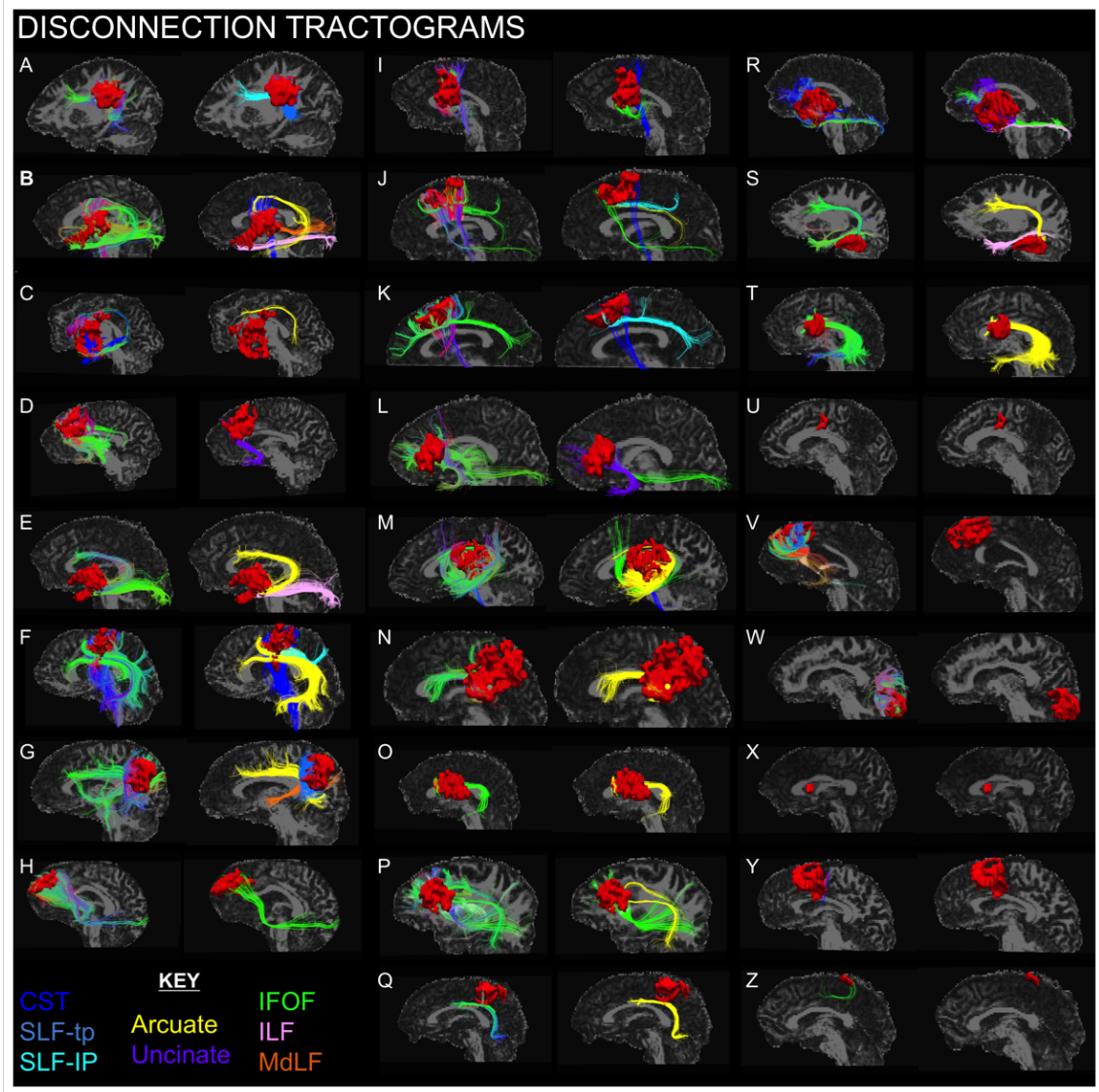


Figure 4.6: Cohort Disconnection Tractogram Output

*Shows the Disconnection Tractograms and identified tracts in this cohort that passed quality control. The left image shows the resection ROI (red) with the raw tractography output, colored by standard direction. The right image shows the tractography output segmented by a human operator to reflect fascicle membership. [CST=cortico-spinal track (dark blue); SLF-tp=superior longitudinal fascicle temporal-parietal component (medium blue); SLF-IP=superior longitudinal fascicle infra-parietal component (light blue); Arcuate=arcuate fascicle (yellow); Uncinate=uncinate fascicle (purple); IFOF=inferior fronto-occipital fascicle (green); ILF=inferior longitudinal fascicle (pink); MdLF=Medial longitudinal fascicle (orange)]*

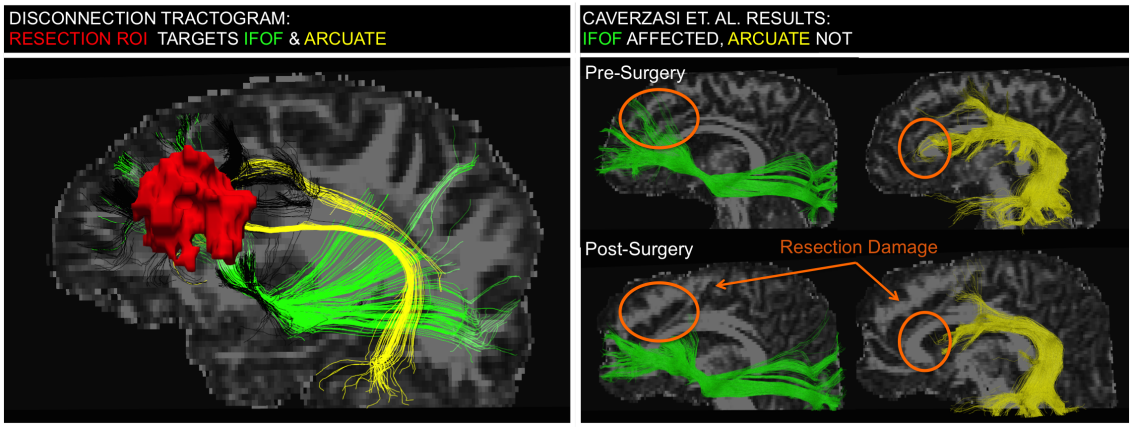


Figure 4.7: Advantage: Continuous Representation of Damage

*From the connectivity of the anisotropic resection ROI (red), portions of the left Arcuate Fascicle (yellow) and left Inferior Fronto-Occipital Fascicle (green) models were segmented. This agrees with the manual reconstructions of the pre- and post-surgery Arcuate and IFOF (RIGHT) models. The IFOF appears to have lost a big chunk to the resection (damage rating 1→2). The Arcuate, however, did not change in subjective rating pre- vs. post-surgery (rating 2→2); the fascicle model was partially destroyed pre-surgery and futher destroyed post-surgery, but the rating system was not resolute enough to reflect that distinction. Manual Results: IFOF (1→2), Uncinate (0→2), SLFIII (2→3, Arcuate (2→2), SLFII (1→1)*

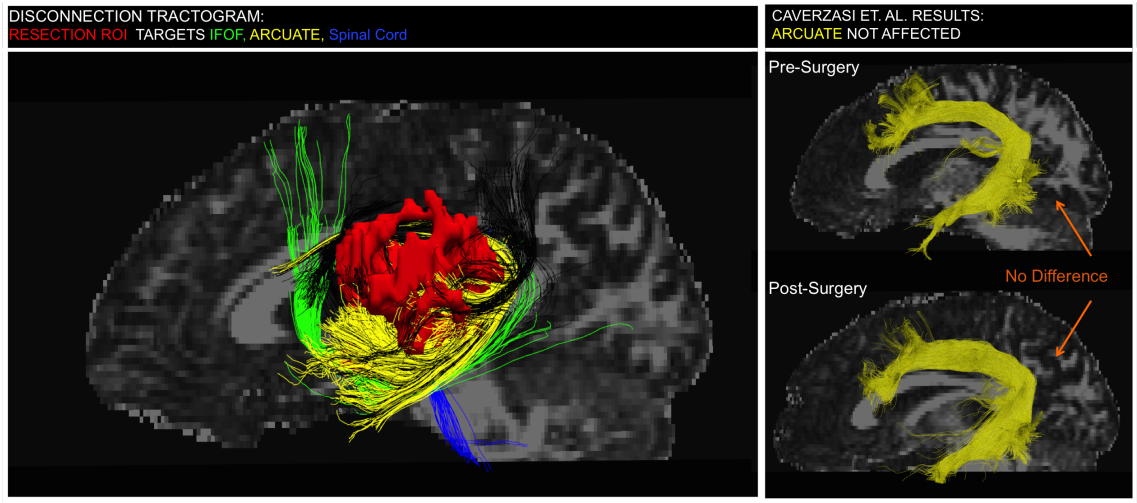


Figure 4.8: Wrapping Around the Tumor: Artifact or Important?

*The region immediately adjacent to the tumor contains tracks that do not meet the criteria for inclusion in the any Anatomically-Constrained fascicle models because both ends do not reach the cortex. However, the bundles follow trajectories characteristic of the Arcuate Fascicle (yellow), the IFOF (green), and a bundle connected to the spinal cord (blue). The pre-surgical model of the Arcuate does not include the component wrapped around the tumor because it does not reach the cortex, so the pre- vs. post-surgical models were rated as unchanged.*

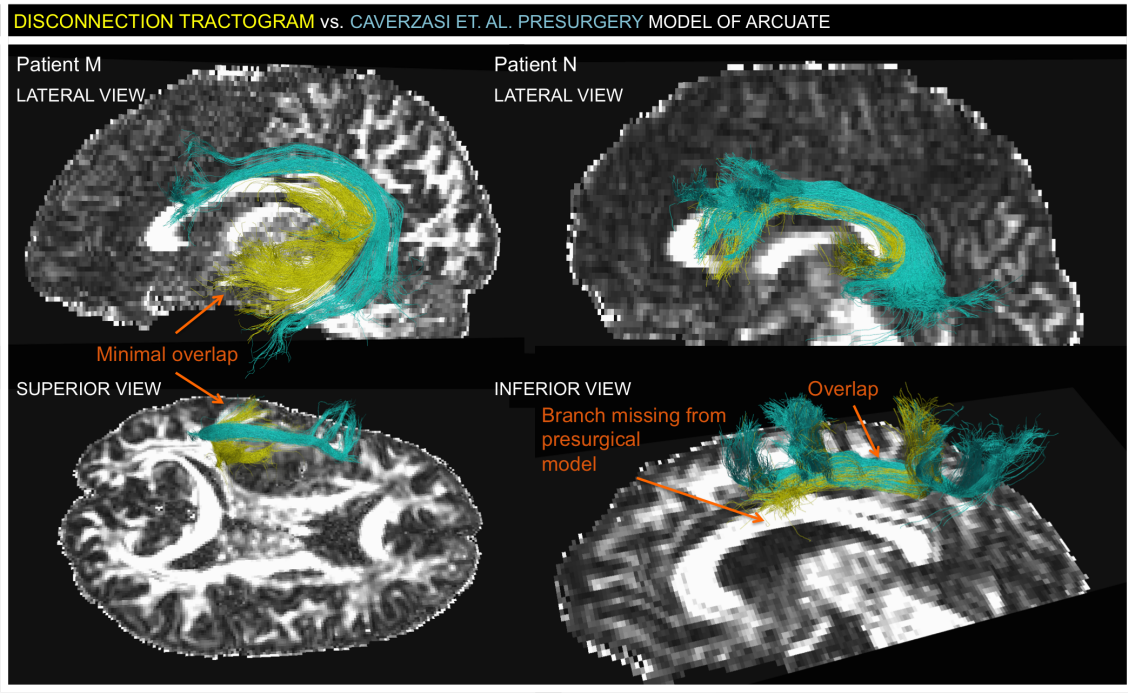


Figure 4.9: Disconnection Tractogram vs. ACT Presurgical Model of Arcuate Fascicle in Patients with Postsurgical Language Deficits

*The presurgical model of the Arcuate Fascicle (Blue) in Patient M (LEFT) is almost entirely distinct from the disconnection tractogram (yellow), which shows a fragment of an Arcuate Fascicle model wrapped around the tumor region. The presurgical model of the Arcuate Fascicle (Blue) in Patient N (RIGHT) largely overlaps with the presurgical model, evidenced by the interdigitation of the blue and yellow streamlines.*

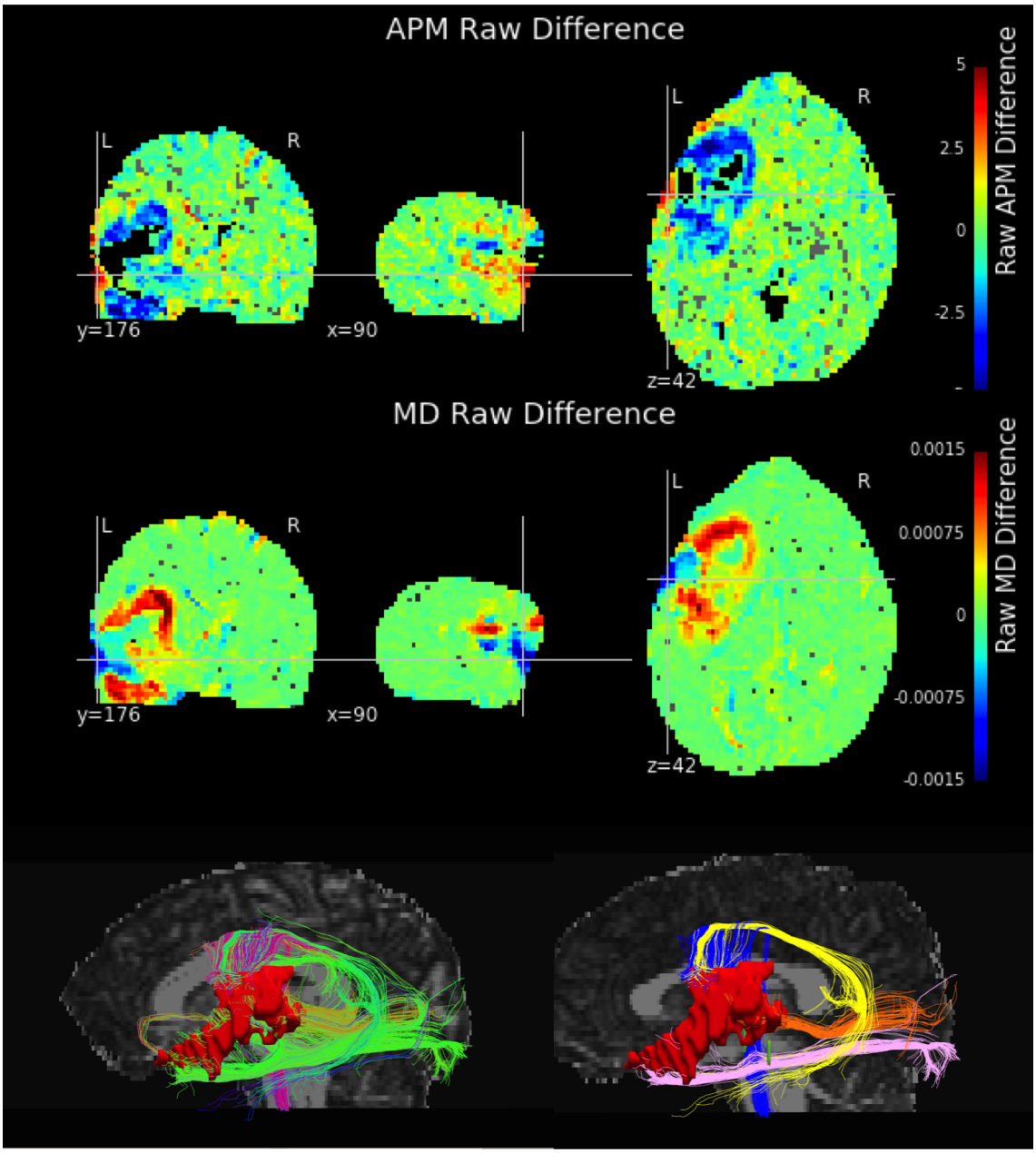


Figure 4.10: Pipeline Performance with Blood Products and Pneumocephalus

*This patient had pneumocephalus and blood products which appears to manifest as inverted contrast on the anisotropic power map (APM) and mean diffusivity (MD). A) Blood and pneumocephalus are apparently associated with hyperintense AP, resulting in an artifactual increase in AP post-surgery. B) Mean diffusivity decreases due to viscous blood, air, and ischemia. C) The difference pipeline still succeeds in isolating the tissue that decreased in AP. This difference map suggsts that the Arcuate (yellow), Medial Longitudinal Fascicle (MdLF=orange), the Inferior Longitudinal Fascicle (ILF), and a bundle connected to the spinal cord (blue) were at risk of being damaged by the surgical resection. Manual Results: Arcuate (1→2), MdLF (1→3), ILF (2→2), SLFII (2→2), SLFIII (3→3), IFOF (3→3), Uncinate (3→3), CST face/hand (1→1)*

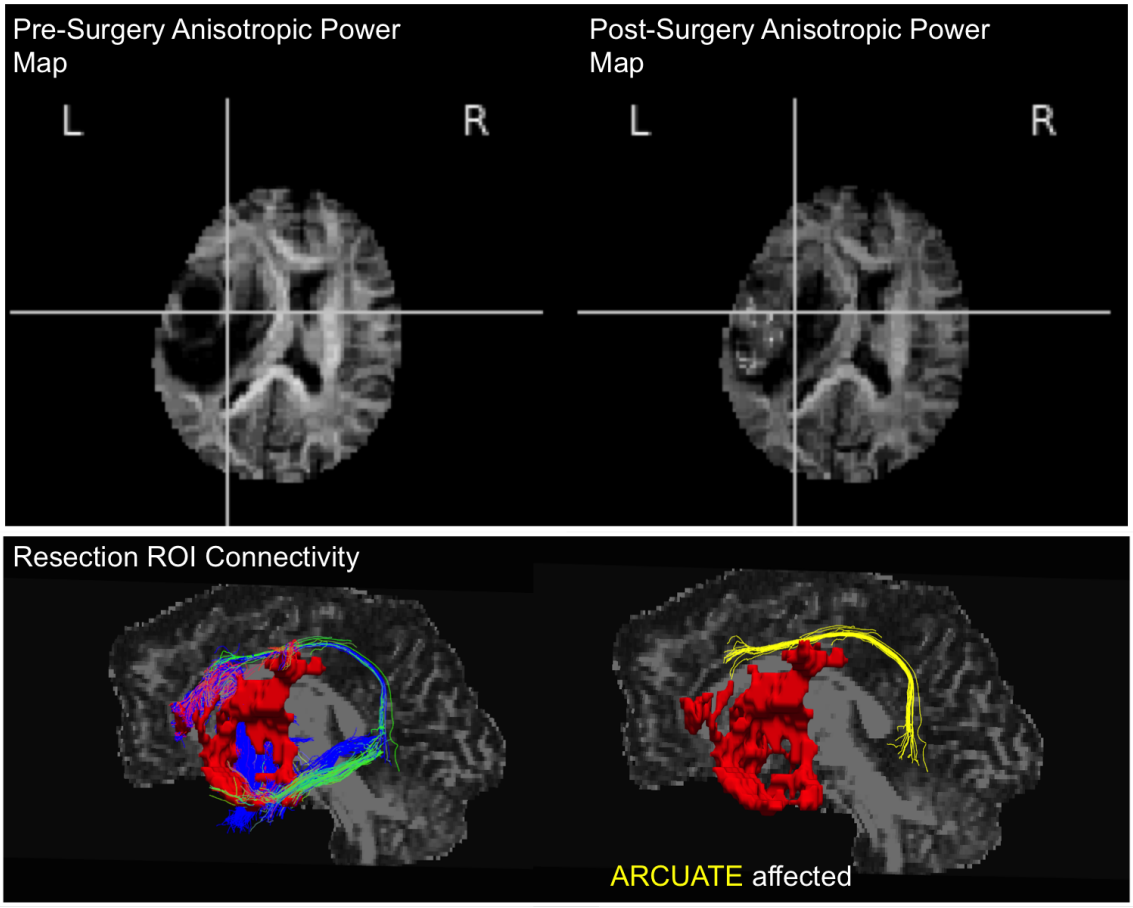


Figure 4.11: Pipeline Performance with Severe Tissue Distortion

*In this case, severe mass effect was observed due to the aggressive growth of the tumor and pressure from hemorrhaging. This pipeline was still able to estimate an anisotropic resection ASAP ROI and produce a prediction that the Arcuate Fascicle was disconnected. The manual rating also cited the SLFII as being damaged, but the underlying white matter is intact so this discrepancy could be explained by post-surgical tracking difficulty. Manual Results: Arcuate (1→2), SLFII (1→2), SLFIII (2→2), SLF-tp (3→3), ILF (3→3), MdLF (3→3), IFOF (3→3), Uncinate (3→3), CST face/hand/foot (1→1)*

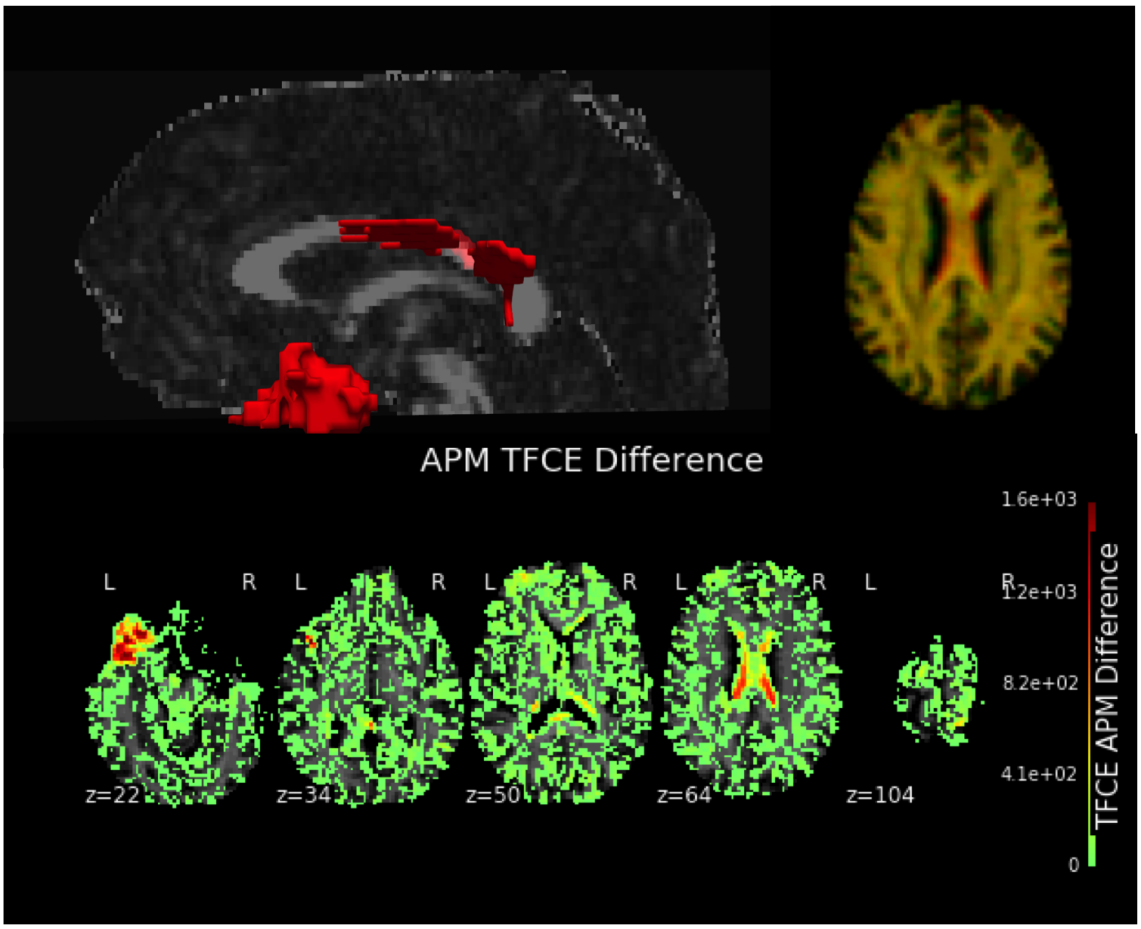


Figure 4.12: Pipeline Performance: False Positive

*In this case, a false-positive error in the ASAP ROI was observed due to excessive distortion of ventricles. Blood products and pneumocephalus appear to have compressed the brain postoperatively and caused the ventricles to shift, in relation to the preoperative image. The peri-ventricular white matter was misclassified as ASAP ROI because parts of the highly anisotropic corpus callosum overlay the shifted ventricle in the post-operative image. This false positive presented in 2 out of 29 patients in the cohort and was addressed easily by QC.*

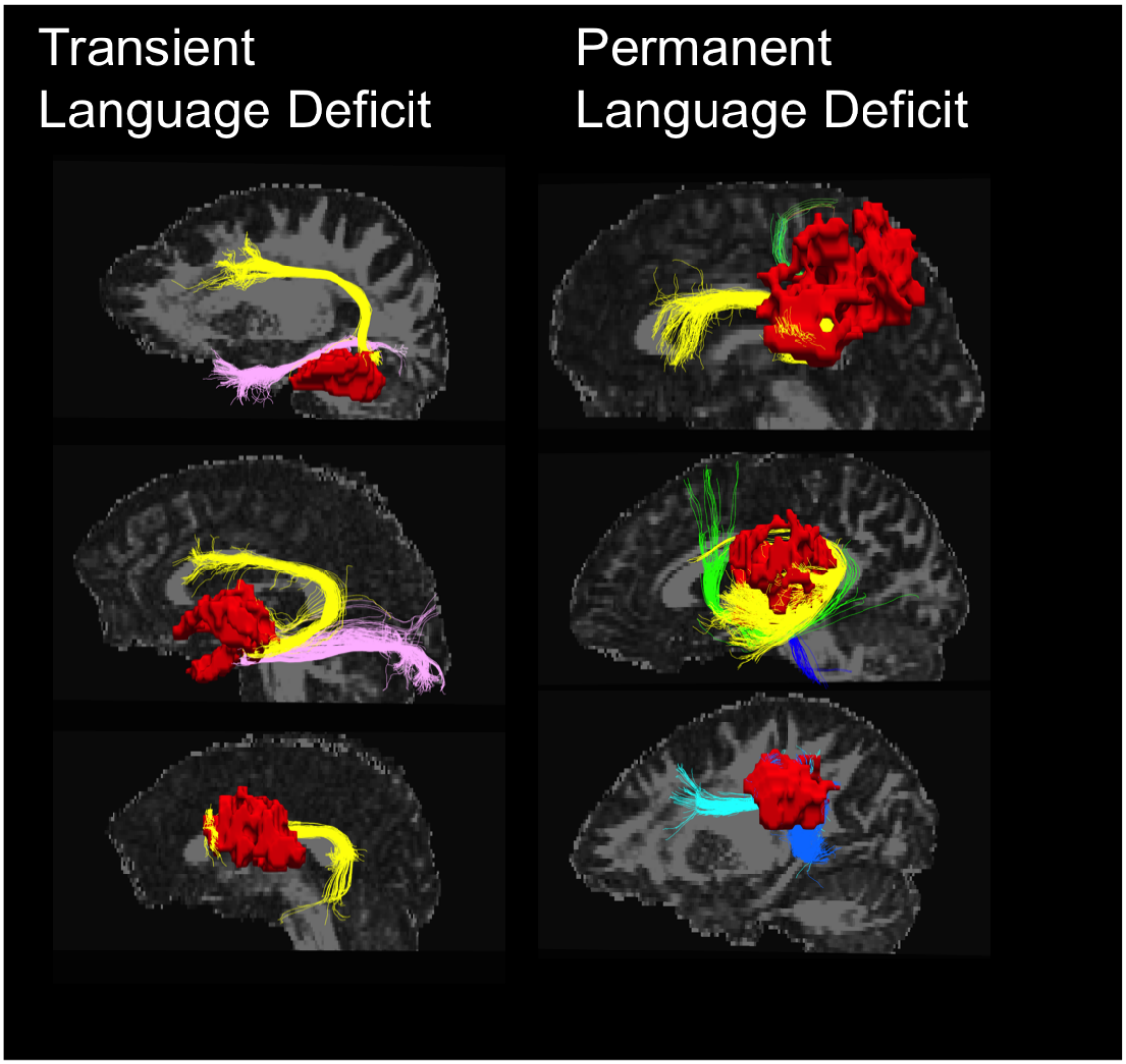


Figure 4.13: Language Deficit Disconnection Tractograms

*Shows the Disconnection Tractograms and identified tracts in this cohort that experienced transient (left) and permanent (right) language deficits, colored by fascicle membership. [CST=cortico-spinal track (dark blue); SLF-tp=superior longitudinal fascicle temporal-parietal component (medium blue); SLF-IP=superior longitudinal fascicle infra-parietal component (light blue); Arcuate=arcuate fascicle (yellow); IFOF=inferior fronto-occipital fascicle (green); ILF=inferior longitudinal fascicle (pink)]*

# Chapter 5

# Tractography-Based Treatment Volume Modification for Radiation Therapy Planning: Pilot Study

## 5.1 Introduction

Glioblastoma multiforme (GBM) is an aggressive and common form of malignant primary brain tumor. The treatment for GBM typically consists of surgery, radiation therapy, and chemotherapy, but the prognosis for patients is still quite poor with a median survival of about 15 months [2]

The standard treatment starts with a maximal safe neurosurgical resection [107]. After surgery, the patient typically receives concurrent radiation therapy (RT) and chemotherapy [151]. Postmortem investigation has suggested that most recurrences (90%) occur within 2cm of the primary tumor site [86], so the RT treatment volume (clinical target volume = CTV) is often defined by a 2cm isotropic expansion around the tumor bed and any residual tumor (Gross Tumor Volume = GTV). It should be noted that there is not a consensus on the optimal way to define the CTV. Several studies have tested the accuracy of different RT volume definitions. Defining the CTV by a 2cm expansion had a success rate of 86% [84]. Another study compared this 2cm expansion CTV to an approach including peritumoral

edema and found that the including edema did not change the failure rate, but did increase the radiated volume [40]. There is evidence that tumor cells migrate preferentially along large white matter bundles [73, 72], so an anisotropic expansion that takes large white matter bundle configuration into account may maximize the chances of radiating migrating cancer cells and minimize the amount of brain tissue radiated. As discussed at length in previous chapters, diffusion-weighted MRI (DW-MRI) can be used in combination with fiber tracking algorithms to model the trajectory of large white matter pathways using the direction and magnitude of water movement in tissue [11, 108, 133, 130, 159, 154, 92, 8, 93, 161, 85]. The method presented here is a tool for translating a diffusion MRI fiber tracking dataset into a pathlength map that assigns each voxel a nearest distance along a streamline back to a provided region of interest.

## 5.2 Methods

### Patients

Twenty-six patients with newly diagnosed Glioblastoma (GBM) who underwent primary adjuvant radiotherapy were selected for this pilot study. Inclusion criteria were: radiographically confirmed recurrence and meeting the standard HARDI scan parameters. Nine patients were excluded because they did not have a radiographically confirmed recurrence and three patients were excluded because their HARDI scans were inconsistent with the cohort. All included subjects underwent a High Angular Resolution Diffusion Imaging scan (HARDI) with the following parameters: 55 directions, b=1000 s/mm2, 1 B0 image, 1.09x1.09mm in-plane resolution, 2mm slice thickness. Research was performed in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by our institution. The Committee on Human Research at the University of California, San Francisco, approved the study protocol. Written informed consent was obtained from all study participants.

### Approach

Given the knowledge that tumor cells preferentially migrate along large white matter bundles, the use of tractography information to model the trajectory of these pathways should increase the chance that a Radiation Therapy (RT) treatment volume includes the site of future recurrence. To incorporate the diffusion tractography information into the RT treatment planning software, a path length map was generated to represent the modeled white matter pathway distance. Each voxel was assigned the value of the minimum distance along a streamline connecting it to the Gross Tumor Volume (GTV), defined as the union of the resection cavity and any residual tumor. A Radiation Therapist hand-drew the GTV to include the resection cavity and any residual disease, using the FLAIR image.

### Open-Source Code

The path length function, implemented in the Diffusion Imaging in Python (DIPY) open-source package [66] (PR 1114) computes the shortest path, along any streamline, between a given region-of-interest (ROI) and each voxel in the provided volume.

Other supporting functions included in this Pull Request (PR 1114: flexi tvis affine, get flexi tvis affine) are used in the path length function to reconcile affine transformations to map between grid and streamline spaces with different voxel orders. These functions are useful in other applications, like targeting streamlines with an ROI and creating renderings with VTK renderings. Otherwise, the user must search through headers and figure out how to get all elements in the same space for image processing or display purposes.

### Radiation Therapy Planning

## 5.3 Results

## 5.4 Discussion

The demonstrated evaluation of a diffusion approach is a good method to quickly evaluate whether or not an anisotropic expansion confers added benefit to standard-of-care isotropic

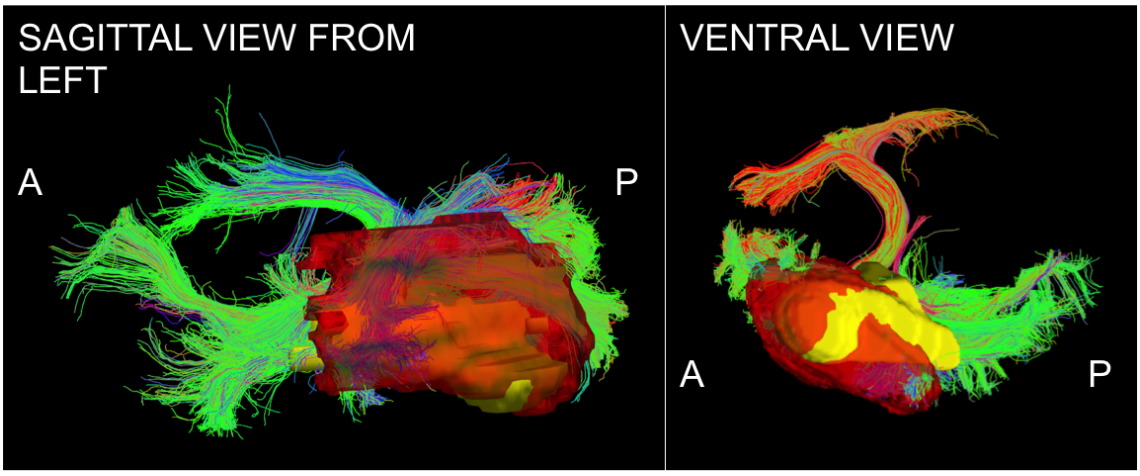


Figure 5.1: Example Patient 1 Fiber Tracking Model of GTV White Matter Connectivity

*The streamline connectivity of the GTV (red volume) is shown with the recurrence ROI (yellow volume)*

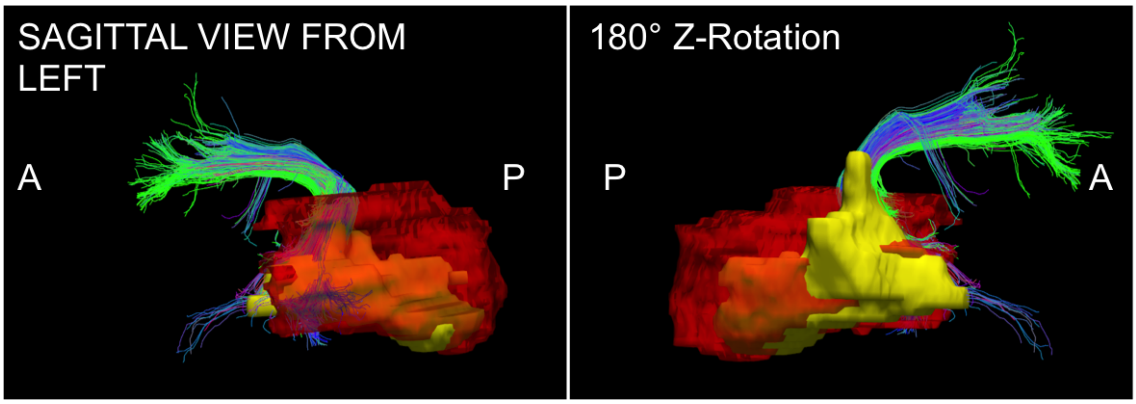


Figure 5.2: Example Patient 1 Recurrence Follows Arcuate Fascicle (DWI15)

*Isolating the Arcuate Fascicle from the streamline connectivity of the GTV (red volume),it is apparent that the recurrence (yellow volume) follows the path of the bundle modeled using tractography*

expansion. If the anisotropic expansion provided by the underlying streamline maps adds value to the prediction of recurrence, the volume of the anisotropic expansion should be less than that of the isotropic recurrence when both are expanded to the recurrence site in a cohort of patients. After added-value has been established, then a much larger study

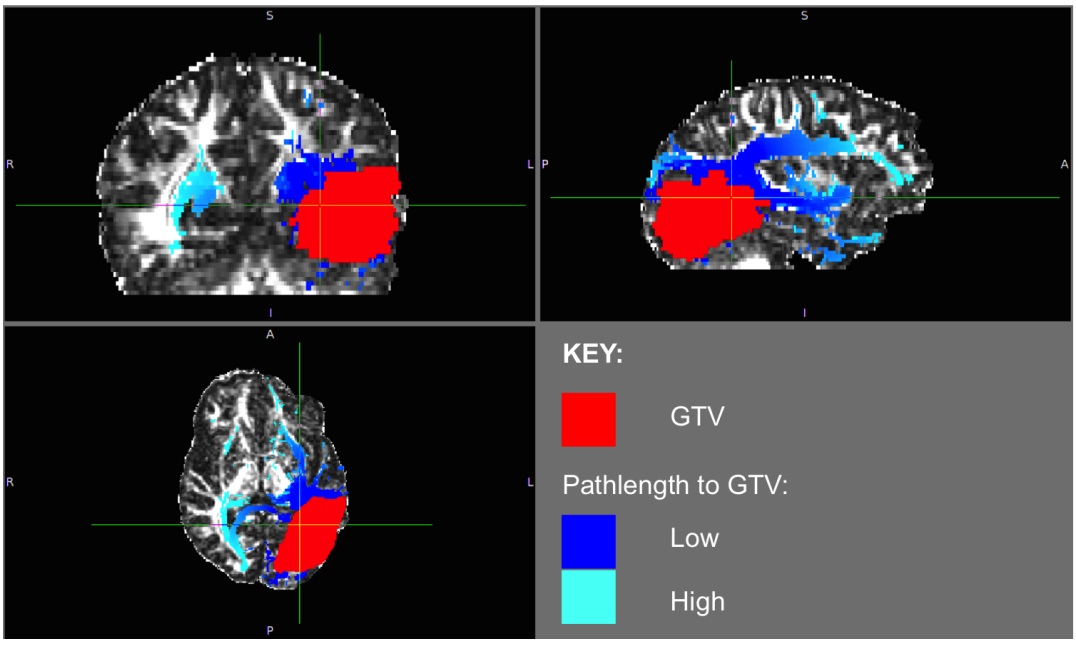


Figure 5.3: Example Patient 1 Pathmap Image from GTV

*The pathmap was generated based on the streamline dataset from the GTV 5.1. The pathmap shows the distance along streamlines from the GTV from low (dark blue) to high (light blue) distance. The GTV is shown in red.*

should be conducted to set expansion parameters so that the methodology can be tested prospectively in a clinical trial.

The flexibility of this approach enables any number of studies to be conducted to investigate the best way to incorporate diffusion information into Radiation Therapy planning. The provided code works on a set of streamlines, so many types of tractography could be investigated, in addition to methods that incorporate human-operator quality-control steps. The target ROI can also be varied to investigate integrating pathway distance from residual tumor, the primary tumor site, or edematous tissue. The timepoint at which the DW-MRI data is collected could also be varied. This pilot was performed using post-surgical pre-radiation therapy HARDI datasets, but pre-surgical HARDI datasets might be more effective to study because the fascicles underlying the tractography have not yet been disturbed by surgical intervention. Many theories of how to use pathway information can be tested with this methodology: for example, large bundles that were cut by the surgery could

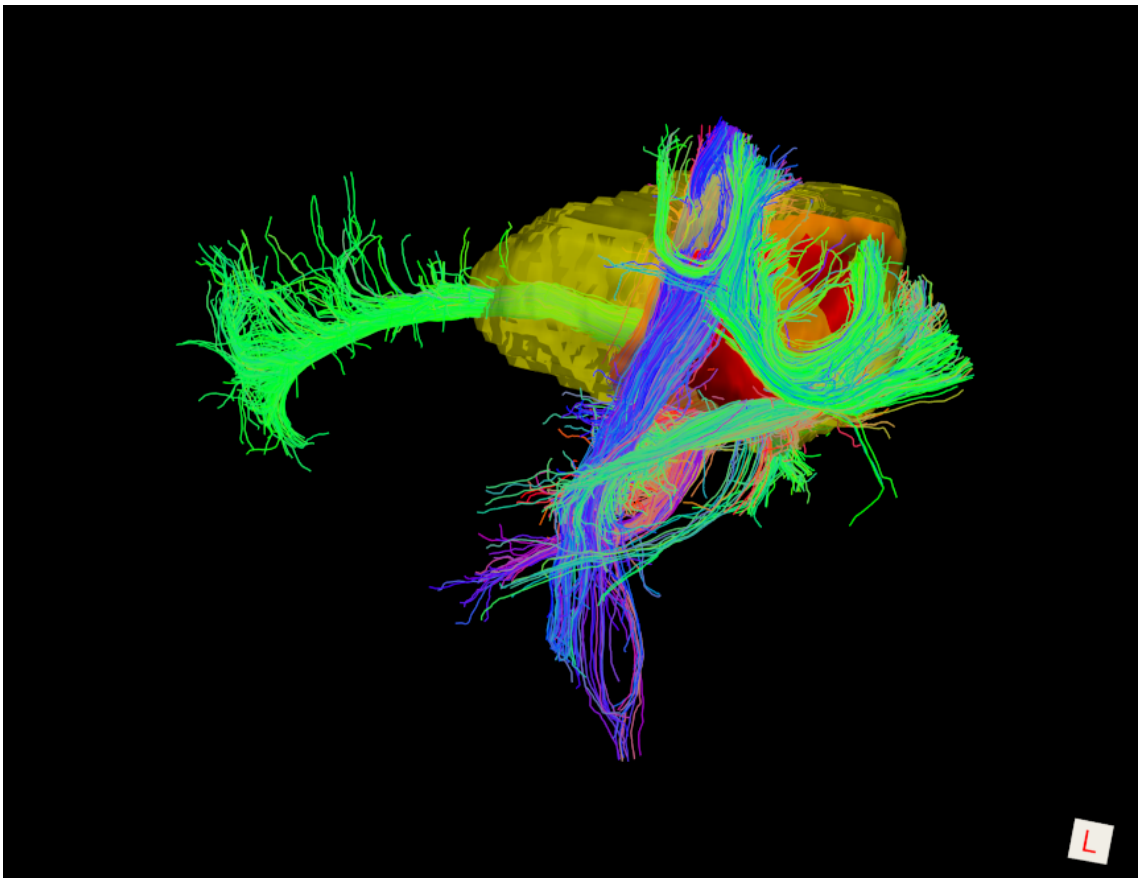


Figure 5.4: Example Patient 2 Fiber Tracking Model of GTV White Matter Connectivity

*The streamline connectivity of the GTV (red volume) is shown with the recurrence ROI (yellow volume), which appears to follow the path of the cingulum bundle*

be isolated to see if neuronal degeneration creates a tumorigenic environment, increasing the chance that the tumor will recur along that pathway. If a set of streamlines can be created to reflect a hypothesis, this framework easily translates the dataset into an added-value analysis to evaluate whether or not that streamline dataset improved prediction of recurrence over isotropic expansion.

These pathmaps can be loaded into a RT planning software to anisotropically modify the treatment volume. A pilot case is shown in Figure 5.6. The pathlength map (orange) was thresholded at a fixed distance (yellow) from the GTV (red) and that volume used to modify the typical isotropic treatment volume (pink).

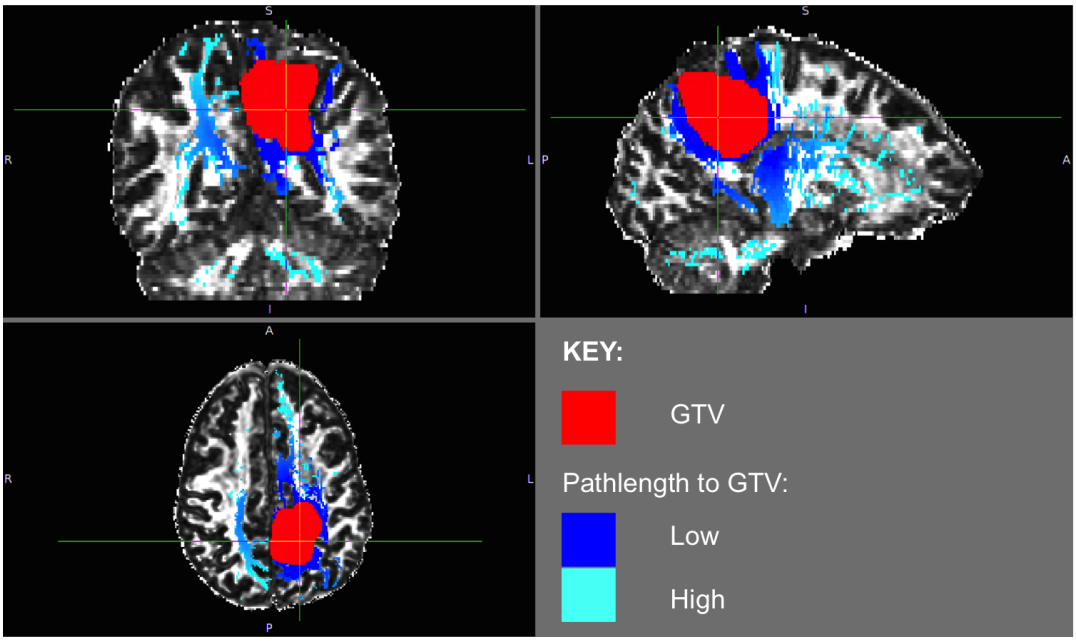


Figure 5.5: Example Patient 2 Pathmap Image from GTV

*The pathmap was generated based on the streamline dataset from the GTV. The pathmap shows the distance along streamlines from the GTV from low (dark blue) to high (light blue) distance. The GTV is shown in red.*

## 5.5 Conclusions

This proof-of-concept demonstrated a tool for integrating diffusion information with existing protocols for Radiation Therapy Planning. The pathlength maps presented in this manuscript can be used to conduct studies predicting tumor recurrence using diffusion data to model large fiber pathways. The code to generate the pathlength maps is open-source and freely available on Github as part of the Diffusion Imaging in Python package [66]. (`https://github.com/nipy/dipy/blob/master/dipy/tracking/utils.py`)

## 5.6 Author Contributions

This work was conducted in collaboration with Michael Wahl, Olivier Morin, Bagrat Amirbekian, Julia Owen, Roland Henry, Steve Braunstein, and Pratik Mukherjee.

Figure 5.6: Example Patient 2 Raystation Anisotropic CTV Based on Pathlength Map

*This is an example of a pathlength map generated based on the streamline dataset from the GTV and imported into Raystation (orange). The GTV (red) is expanded anisotropically using the pathlength map (yellow), which follows the path of the recurrence better than the typical isotropic expansion treatment volume (pink) Image credit: Olivier Morin*

## 5.7 Funding

This work was supported by a UCSF Joint Radiation Oncology/Radiology Seed Grant. KJ was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program and the National Institutes of Health [5R01NS066654-05].

## 5.8 Acknowledgments

Thank you to the Neuroimaging in Python (NIPY) development community, especially the Diffusion Imaging in Python (DIPY) open-source developers, and all of our research subjects at the University of California, San Francisco.

# Chapter 6

# White Matter Connectivity of Cortical Stimulation-Induced Mood Improvement: Pilot Study

## 6.1 Introduction

Direct electrical stimulation of targeted brain regions has been highly successful in the treatment of movement disorders, like Parkinsons Disease [129, 120]. To treat a patient with Deep Brain Stimulation (DBS), a surgeon implants electrodes in specific brain regions so that electrical current can be delivered as a treatment to alter the function of brain circuitry. In addition to certain movement disorders, DBS has been approved by the Food and Drug Administration (FDA) to treat obsessive-compulsive disorder and epilepsy [59, 140]. Many DBS implantation sites have been tested to address treatment-resistant depression [114, 26, 19, 142, 96, 144], but a series of clinical trial failures [143] suggest that the mechanisms behind the anti-depressive therapeutic effect of DBS needs to be better understood. DBS of the subcallosal cingulate has been studied extensively [136, 47] as a therapy for treatment-resistant depression. This approach has evolved from anatomical targeting of a region associated with positron emission tomography (PET) changes in successful treatment response to a tractography-based approach that targets the intersection of four white matter bundles:

the forceps minor, uncinate fasciculus, cingulum, and fronto-striatal connections. This DBS site has been tested in a cohort of 77 patients implanted at eight different institutions with highly variable success rates (between 33 and 87.5%) [115, 137, 135, 47, 136]. The potential for tractography-assisted planning of DBS treatment of psychiatric disorders is promising, but limitations of tractography methodology (discussed at length preceding sections of this dissertation) are equally true in this field and must be taken into consideration [122].

A study conducted at the University of California, San Francisco [146] examined a cohort of twelve patients undergoing intracranial monitoring for planning of their epilepsy treatment. These patients volunteered to undergo a stimulation study in which their intracranial electrodes were used to deliver current to the brain while measurements of their mood were taken, under the supervision of an Epileptologist. This gave researchers the opportunity to test different sites for efficacy in improving mood. The results suggested that the lateral Orbito-Frontal Cortex (OFC) is a promising target as a therapy for mood disorders [146]. Prefrontal regions, including the lateral OFC, are thought to mediate mood through a top down mediation of reward pathways [139]. Surgical lesions of the OFC and Cingulate Cortex have been associated with emotional changes, measured subjectively [88]. This chapter analyzes the white matter connectivity of the electrical stimulation sites that elicited a mood improvement in order to generate hypotheses of the brain circuitry that is being disrupted by lateral OFC stimulation in this cohort.

## 6.2 Methods

### Subjects

A cohort of twelve patients that were undergoing inpatient intracranial monitoring for seizure localization as part of their intractable epilepsy clinical evaluation at the University of California, San Francisco volunteered to participate in this study. Patients were implanted with a combination of subdural (high-density grid: AdTech 256 channels, 4mm center-to-center spacing and 1.17mm diameter; strip: 1cm spacing) and depth electrodes (5mm spacing) as part of their clinical evaluation. If sufficient data was collected from the intracranial electrodes prior to the scheduled surgery, the patients underwent a stimulation protocol under

the supervision of an Epileptologist as part of this study. Research was performed in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) and the standards established by our institution. The Committee on Human Research at the University of California, San Francisco, approved the study protocol. Written informed consent was obtained from all study participants.

## Stimulation Protocol

Bipolar stimulation (100 Hz, 100 usec pulse width, 1-6 mA intensity) was applied to electrode pairs with durations of 1 second to continuous for 2-3 minutes. Patients were monitored during both treatment and sham conditions to quantify their mood using two self-reported metrics: Immediate Mood Scaler (IMS) and word valence. During the IMS task, the patient answered questions about their mood on a touch pad. For the word valence task, the patient engaged in conversation with an examiner. An audio recording of the conversation was evaluated for positive or negative word content. These scores were normalized to a 100 point scale and averaged to produce a composite mood score. One patient was not able to complete the IMS task because they were not fluent in English, so their score is entirely composed of word valence. Response was calculated as a percentage increase in composite mood score in the stimulation condition over baseline (the sham condition). The six patients with the highest composite score "response all increased over 30% to baseline, so these patients were analyzed separately to explore white matter connectivity common to stimulation sites eliciting a treatment-like effect of mood improvement.

## Imaging

Patients underwent Magnetic Resonance Imaging (MRI) prior to implantation surgery on a 3T General Electric Medical Systems scanner (Discovery MR750). Datasets acquired included: a High Angular Resolution Diffusion Imaging sequence (TR/TE =6425/80 ms, 50 axial slices, 2.2 $mm^2$ in-plane resolution (interpolated to 1.1 $mm^2$), 2mm slice thickness, b-value 2000 $s/mm^2$, 55 diffusion gradients, 1 minimally diffusion weighted image) and a 3D high-resolution T1-weighted IRSPGR (TR/TE/TI =7/2/400 ms, 180 axial slices 1 mm

thickness, 0.94 x 0.94 $mm^2$ in plane resolution). Post-implantation of intracranial electrodes, patients underwent a Computed Tomography (CT) scan to localize electrodes.

## Image Processing

### Electrode Localization to T1 Space

The electrode localization was performed using the semi-automated anatomical labeling and inter-subject warping pipeline developed in-house by the Chang Lab at UCSF, published as an open-source python package [78]. Briefly: electrode locations were selected manually on the CT scan and transformed to the T1 space by a rigid registration, followed by a projection onto a hull of the pial surface to correct for brain compression caused by grid placement.

### Diffusion Preprocessing

The datasets were first corrected for motion and eddy current distortion using the FMRIB Software Library (FSL) [91] and the gradient table rotated [110]. A tensor model was fit to the corrected HARDI data using the open-source package Diffusion Imaging in Python (Dipy) [66] and the resulting parameters used to calculate tensor metrics. An Anisotropic Power Map was calculated (APM) [49] using the Q-ball model [50] implemented in Dipy. This diffusion-derived image can be used to perform a better registration to the T1 space than FA, and has been proposed as a solution to susceptibility correction in datasets that have not acquired a fieldmap or "blip-up-blip-down" B0 image [41]. The B0 image was isolated from the HARDI series and used to create a brain mask by skull stripping the image with FSL's Brain Extraction Tool (BET) [149]. This brain mask was applied to all images.

### Tractography Modeling of White Matter Connectivity

Whole-brain residual bootstrap probabilistic q-ball tractography [23] was performed with the parameters described in [37]. Briefly, the preprocessed diffusion signal was fit using spherical harmonics (even orders up to 4) so that orientation distribution functions (ODFs) with constant solid angle factor could be estimated from the data [157, 155, 156]. To perform residual-bootstrap tractography [23], principal fiber orientation was estimated at each step by

computing a bootstrapped ODF and identifying the peaks. The principal fiber orientations from the bootstrap ODFs provided the distribution of fiber tracking directions. Tracking was seeded uniformly at a density of $3^3$ seeds per voxel and terminated by the following criteria: APM threshold of 4 [49] and maximum angle of 60 [29].

### Electrode Targeting

For each patient, a diffeomorphic registration was performed to align the anisotropic power map (APM) and the T1 image to transfer the electrodes from the T1 image to diffusion space using the Advanced Normalization Tools (ANTs) symmetric diffeomorphic registration [10, 9], implemented in DIPY [66]. The dataset did not include images for appropriate susceptibility correction, so the registrations and the point transformation to diffusion space was manually quality-controlled by comparison to cortical anatomy. Trackvis was used to visualize results [163] and target the whole-brain tractography. Using python software written in-house, the white matter voxel nearest to the centroid between the positive and negative bipolar electrodes was identified. Using Trackvis, a human operator expanded a sphere centered on this centroid white matter voxel until the first streamline was targeted. This point in space was selected as the centroid of a uniform 5mm sphere used to target the whole-brain tractography datasets in each patient. This procedure ensured that the geometry of the tissue did not penalize cortical electrodes; tractography representing the white matter connectivity of each electrode was targeted 5mm deep into "trackable" white matter so that the most likely white matter connections serving the cortical regions was represented in each patient. Any streamlines less than 40mm in length were excluded.

### MNI Heat-map Generation

To generate the group results on an MNI template, a diffeomorphic registration was performed between the anisotropic power map (APM) and MNI template using the Advanced Normalization Tools (ANTs) symmetric diffeomorphic registration [10, 9], implemented in DIPY [66]. The streamline dataset representing the connectivity of each patient was binarized such that each voxel containing at least one streamline was included in the binary mask. The diffeomorphic diffusion-to-MNI transformation was applied to the binarized masks so

that they could be summed in the common MNI space to produce heat maps that represent the spatial overlap between white matter connectivity in multiple patients. This analysis was performed to show the connectivity of the entire cohort with the right (orange-yellow) and left (blue-green) hemispheric electrodes summed independently in Figure 6.1, and to show the white matter connectivity common to electrodes that elicited a treatment-like response (Figure 6.2)

## 6.3 Results

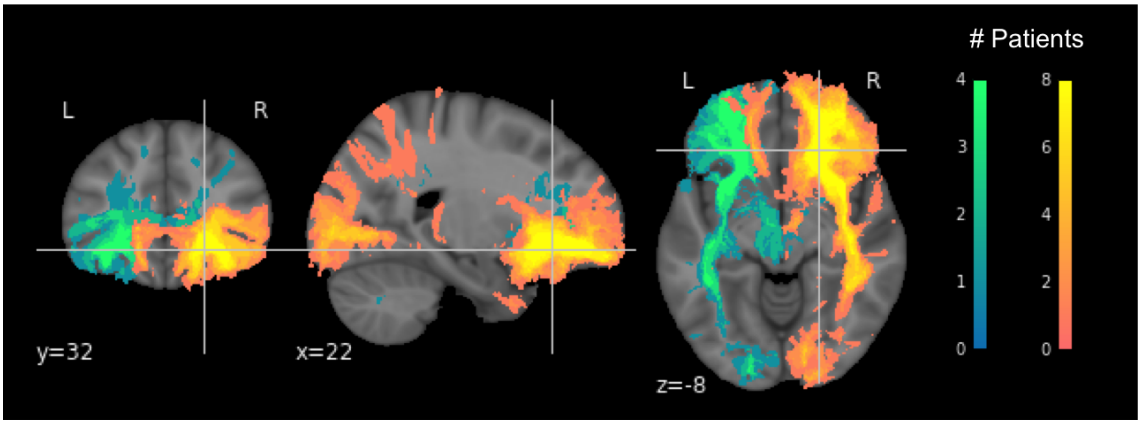


Figure 6.1: Tractography Connectivity Model from Lateral OFC Stimulation in All Patients

*This is a heatmap of the white matter connectivity of the lateral OFC stimulation electrodes across all twelve patients in the cohort. Eight patients were stimulated in the right hemisphere (orange-yellow) and four patients were stimulated in the left hemisphere (blue-green)*

## 6.4 Discussion

The white matter connectivity common to patients exhibiting mood responses of greater than 30% from baseline appears to follow the trajectory of the Frontal Orbito-Polar Track (FOP), a white matter bundle that connects the posterior and anterior orbitofrontal gyri (pOFC and aOFC) to the inferior polar cortex [35]. Tractography representations of the FOP have been published by [35] with validation through blunt dissection, supported by similar connections

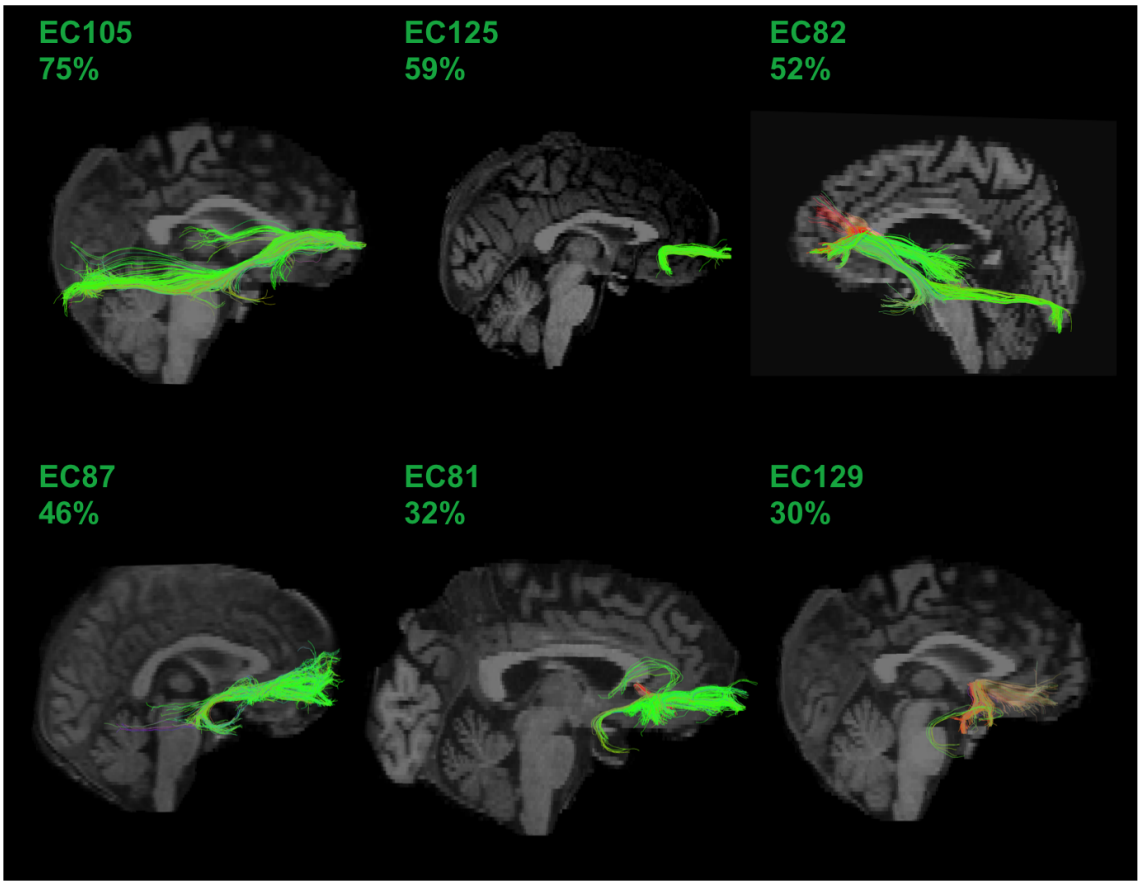


Figure 6.2: Tractography Connectivity Model from Mood Improving Lateral OFC Stimulation

*This is the streamline connectivity of the six most responsive patients. Each patient identification is specified by EC and the percentage improvement in composite mood score over baseline is noted for each case.*

found in monkey brains [134, 170, 145]. In a macaque tracer study of the ventral prefrontal cortex (vPFC) conducted by [113], the authors defined connections of the Uncinate Fasciculus (UF) to include three components: vPFC to temporal cortex connections, vPFC to vPFC connections, and bundles that provide a "conduit" joining vPFC neurons to other tracks. Fractional anisotropy and volume of uncinate fascicle models were associated with social anxiety disorder in a study by [13] and trait anxiety predicted the fractional anisotropy in pathways between the amygdala and the ventromedial PFC and OFC in a study by [58]. In a study by [14], functional connectivity between the anterior insula and the basolateral

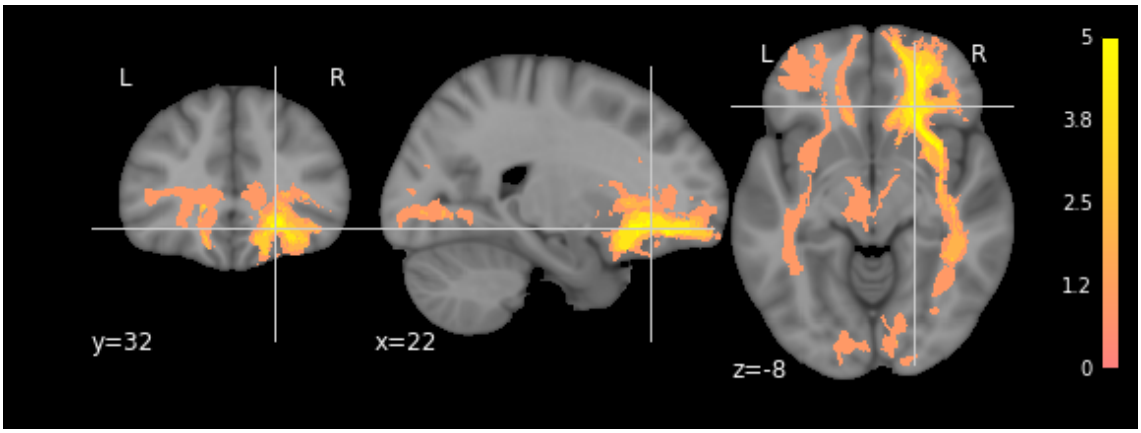


Figure 6.3: Tractography Connectivity Model from Successful Lateral OFC Stimulation: Heatmap

*This is a heatmap of the white matter connectivity of the lateral OFC stimulation electrodes across the six most responsive patients. Five of the highest responding patients were stimulated in the right hemisphere and one was stimulated in the left hemisphere. The "responding" Direct Electrical Stimulation Sites tended to be connected to the frontal pole, according to the tractography analysis.*

amygdala were associated with state anxiety (explained 40% between-subject variance) and axial diffusivity was associated with trait anxiety in the pathway modeled using tractography between the same two regions. While implementing the methods detailed by [35] to model the FOP, the authors noted extensive colocalization of streamlines in the anterior-posterior frontal lobe projections of the FOP and UF that made the FOP difficult to segment in some patients. Given this colocalization, it is possible that results reflecting changes to the FOP could be attributed to the UF in studies that do not account for the bundles individually. Further study is needed to tease apart the contributions of these pathways to predicting mood disorder metrics.

Examining these results in the context of a series of studies from the Mayburg group [115, 122, 137, 135, 47, 137, 136] presents an opportunity to compare the results from two very different approaches that elicited positive treatment-like effects. These sequential studies from the Mayburg group compared the white matter connectivity, modeled using tractography, in patients that had a treatment response and those that did not. The target brain region of the Mayburg studies is the subcallosal cingulate white matter (SCC), which

they define using tractography to identify the intersection between four white matter bundles: forceps minor, uncinate fasciculus, cingulum and fronto-striatal connections.

The results from our current study described in this manuscript elicited a treatment-like response by stimulating cortically in the lateral orbitofrontal cortex. The white matter connectivity common to these studies is the frontal pole; the Mayburg target reaches the frontal pole through the forceps minor and the uncinate, and the study presented in this manuscript connects to the frontal pole via the FOP. In the study by [135], the patients that never responded to stimulation treatment during the two year study were strikingly missing connectivity to the fontal pole. Additional support for the importance of frontal pole stimulation in this study comes from the six patients that converted from nonresponders to responders between the six month and two year timepoints: the tractography connectivity maps changed to include more medial frontal pathways going to the frontal pole [135].

Several limitations of this pilot study must be kept in mind. The cohort is small, and the electrodes are split between two hemispheres (8 Right; 4 Left). Only patients with a stimulation-related mood improvement of 30% over baseline were included in the treatment-effect hypothesis generation analysis (Figures 2 and 3), which effectively halves the sample size to six, five of which are in the right hemisphere. There are also limitations associated with the mood quantification tasks (self-reported IMS and Word Valence), which are variable and challenging to interpret. The lack of a field map or "blip-up-blip-down" B0 images to correct for susceptibility is another major limitation of this study. The authors strongly recommend acquiring one of these images to enable susceptibility correction using a tool like those provided in the FMRIB Software Library (FSL) [6, 150]. All results presented in this manuscript must be considered with the caveat that susceptibility distortions were acting on the diffusion data, which is particularly severe in the OFC due to the air-tissue interface with the sinus. To address these concerns, a registration was performed between an anisotropic power map derived from the HARDI data [49] and the T1 image, as described in [41].

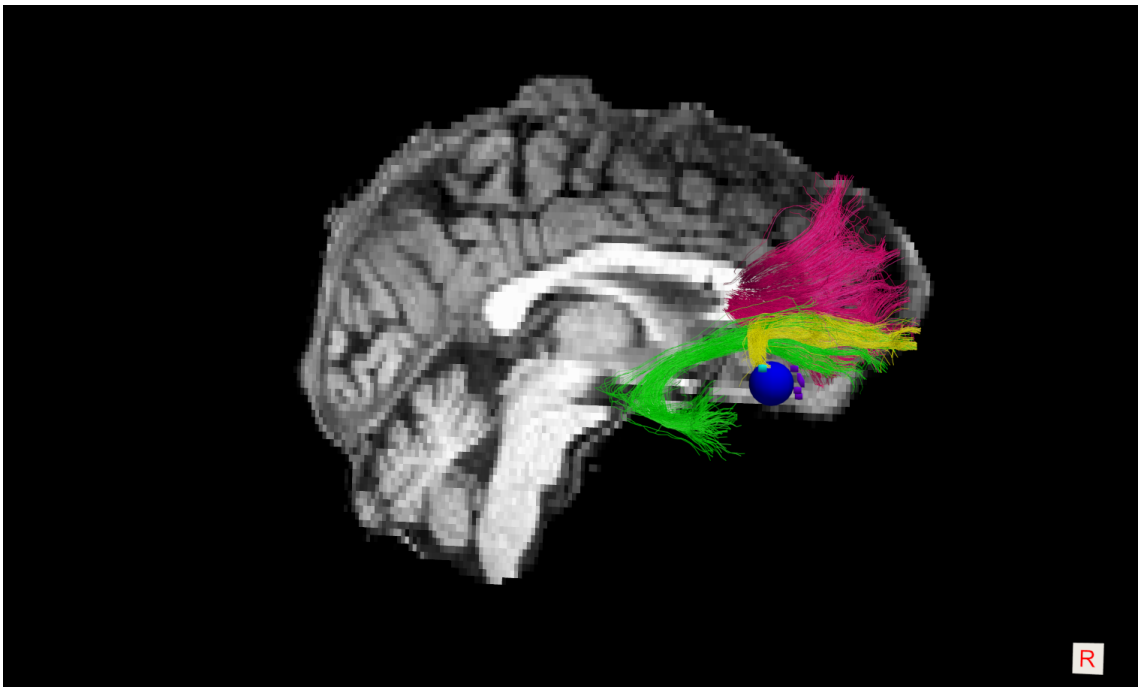


Figure 6.4: Anatomically Constrained Tractography of Frontal Orbito-Polar, Uncinate, and Forceps Bundles

*Tractography models of the Frontal Orbito-Polar, Uncinate, and Forceps bundles are shown in subject EC125*

## 6.5 Conclusions

This pilot project lays the groundwork for hypothesis-driven study of the white matter underlying mood improvement effects elicited by direct electrical stimulation of the lateral Orbitofrontal Cortex. The demonstrated pipeline addresses the challenges associated with tractography in the ventral frontal lobe to the highest degree possible with the available data, but the limitations must be kept in mind because modern susceptibility correction could not be executed. The results of this study generate the hypothesis that the successful mood improvement associated with lateral OFC stimulation may be acting through a connection to the frontal pole. Future studies examining the anatomical as well as functional connections of the lateral OFC to the frontal pole via the Frontal Orbito-Polar Track white matter bundle are necessary to further substantiate this hypothesis. Stimulation of the lateral OFC and engagement of the Frontal Orbito-Polar Track may be novel targets in alleviation of mood

disorder symptoms.

## 6.6 Author Contributions

This work was conducted in collaboration with Lucia Bederson, Deanna Wallace-Donovan, Kristin Sellers, Vikram Rao, Liberty Hamilton, David Chang, Morgan Lee, Heather Dawes, Edward Chang, Heather Dawes, and Roland Henry.

## 6.7 Funding

This research was partially funded by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement Number W911NF-14-2-0043, issued by the Army Research Office contracting office in support of DARPAS SUBNETS program. The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. KJ was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program and the National Institutes of Health [5R01NS066654-05].

## 6.8 Acknowledgments

Thank you to the Neuroimaging in Python (NIPY) development community, especially the Diffusion Imaging in Python (DIPY) open-source developers, and all of our research subjects.

# Bibliography


[1] I. Aganj, C. Lenglet, and G. Sapiro. "Accurate ODF Reconstruction in Q-ball Imaging". In: *NeuroImage* 47 (July 2009), S51.

[2] C. Alifieris and D. T. Trafalis. "Glioblastoma multiforme: Pathogenesis and treatment". In: *Pharmacol. Ther.* 152 (Aug. 2015), pp. 63–82.

[3] Fabien Almairac et al. "The left inferior fronto-occipital fasciculus subserves language semantics: a multilevel lesion study". In: *Brain Structure and Function* 220.4 (Apr. 2014), pp. 1983–1995.

[4] Bagrat Amirbekian, Ariel Rokem, and Elefarios Garyfalladis. *An introduction to the Probabilistic Direction Getter*. `https://github.com/nipy/dipy/blob/master/doc/examples/probabilistic_fiber_tracking.py`. 2016.

[5] Sarah M. Andersen, Steven Z. Rapcsak, and Pélagie M. Beeson. "Cost function masking during normalization of brains with focal lesions: Still a necessity?" In: *NeuroImage* 53.1 (Oct. 2010), pp. 78–84.

[6] Jesper L.R. Andersson, Stefan Skare, and John Ashburner. "How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging". In: *NeuroImage* 20.2 (Oct. 2003), pp. 870–888.

[7] John Ashburner and Karl J. Friston. "Unified segmentation". In: *NeuroImage* 26.3 (July 2005), pp. 839–851.

[8] Yaniv Assaf and Peter J. Basser. "Composite hindered and restricted model of diffusion (CHARMED) MR imaging of the human brain". In: *NeuroImage* 27.1 (Aug. 2005), pp. 48–58.

[9] Brian B. Avants et al. “A reproducible evaluation of ANTs similarity metric performance in brain image registration”. In: *NeuroImage* 54.3 (Feb. 2011), pp. 2033–2044.

[10] B AVANTS et al. “Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain”. In: *Medical Image Analysis* 12.1 (Feb. 2008), pp. 26–41.

[11] P.J. Basser, J. Mattiello, and D. Lebihan. “Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo”. In: *Journal of Magnetic Resonance Series B* 103.3 (Mar. 1994), pp. 247–254.

[12] Elizabeth Bates et al. “Voxel-based lesion–symptom mapping”. In: *Nature Neuroscience* (Apr. 2003).

[13] Volker Baur et al. “Evidence of frontotemporal structural hypoconnectivity in social anxiety disorder: A quantitative fiber tractography study”. In: *Human Brain Mapping* 34.2 (Nov. 2011), pp. 437–446.

[14] Volker Baur et al. “Resting-State Functional and Structural Connectivity Within an Insula–Amygdala Route Specifically Index State and Trait Anxiety”. In: *Biological Psychiatry* 73.1 (Jan. 2013), pp. 85–92.

[15] C. Beaulieu. “The Biological Basis of Diffusion Tractography”. In: *3rd IEEE International Symposium on Biomedical Imaging: Macro to Nano 2006.* Institute of Electrical and Electronics Engineers (IEEE), 2006.

[16] T.E.J. Behrens et al. “Characterization and propagation of uncertainty in diffusion-weighted MR imaging”. In: *Magnetic Resonance in Medicine* 50.5 (Oct. 2003), pp. 1077–1088.

[17] T.E.J. Behrens et al. “Probabilistic diffusion tractography with multiple fibre orientations: What can we gain?” In: *NeuroImage* 34.1 (Jan. 2007), pp. 144–155.

[18] Mohammed Benzagmout, Peggy Gatignol, and Hugues Duffau. “Resection of World Health Organization Grade II Gliomas Involving Broca’s Area”. In: *Neurosurgery* 61.4 (Oct. 2007), pp. 741–753.

[19] Isidoor O. Bergfeld et al. “Deep Brain Stimulation of the Ventral Anterior Limb of the Internal Capsule for Treatment-Resistant Depression”. In: *JAMA Psychiatry* 73.5 (May 2016), p. 456.

[20] Jeffrey I. Berman. “Advanced Diffusion MR Tractography for Surgical Planning”. In: *Functional Brain Tumor Imaging*. Springer Nature, Sept. 2013, pp. 183–194.

[21] Jeffrey I. Berman et al. “Accuracy of diffusion tensor magnetic resonance imaging tractography assessed using intraoperative subcortical stimulation mapping and magnetic source imaging”. In: *Journal of Neurosurgery* 107.3 (Sept. 2007), pp. 488–494.

[22] Jeffrey I Berman et al. “Diffusion-tensor imaging-guided tracking of fibers of the pyramidal tract combined with intraoperative cortical stimulation mapping in patients with gliomas”. In: *Journal of neurosurgery* 101.1 (2004), pp. 66–72.

[23] Jeffrey I. Berman et al. “Probabilistic streamline q-ball tractography using the residual bootstrap”. In: *NeuroImage* 39.1 (Jan. 2008), pp. 215–222.

[24] B. Bernal and A. Ardila. “The role of the arcuate fasciculus in conduction aphasia”. In: *Brain* 132.9 (Aug. 2009), pp. 2309–2316.

[25] Anna Berti, Francesca Garbarini, and Marco Neppi-Modona. “Disorders of Higher Cortical Function”. In: *Neurobiology of Brain Disorders*. Elsevier BV, 2015, pp. 525–541.

[26] Bettina H. Bewernick et al. “Nucleus Accumbens Deep Brain Stimulation Decreases Ratings of Depression and Anxiety in Treatment-Resistant Depression”. In: *Biological Psychiatry* 67.2 (Jan. 2010), pp. 110–116.

[27] Nicole M. Petrovich Brennan and Andrei I. Holodny. “Use of Advanced Neuroimaging (fMRI DTI/Tractography) in the Treatment of Malignant Gliomas”. In: *Malignant Brain Tumors*. Springer Nature, Dec. 2016, pp. 3–13.

[28] Matthew Brett, Alex Leff, and John Ashburner. “Automated nonlinear coregistration of damaged brains to a normal template using cost function masking”. In: *NeuroImage* 11.5 (May 2000), S566.

[29] Monica Bucci et al. "Quantifying diffusion MRI tractography of the corticospinal tract in brain tumors with deterministic and probabilistic methods". In: *NeuroImage: Clinical* 3 (2013), pp. 361–368.

[30] Hugh W. Buckingham. "Semantic Paraphasia". In: *Encyclopedia of Clinical Neuropsychology*. Ed. by Jeffrey S. Kreutzer, John DeLuca, and Bruce Caplan. New York, NY: Springer New York, 2011, pp. 2248–2250. ISBN: 978-0-387-79948-3.

[31] Serena Campana, Carlo Caltagirone, and Paola Marangolo. "Combining Voxel-based Lesion-symptom Mapping (VLSM) With A-tDCS Language Treatment: Predicting Outcome of Recovery in Nonfluent Chronic Aphasia". In: *Brain Stimulation* 8.4 (July 2015), pp. 769–776.

[32] M. Catani. "The rises and falls of disconnection syndromes". In: *Brain* 128.10 (July 2005), pp. 2224–2239.

[33] M Catani and M Thiebautdeschotten. "A diffusion tensor imaging tractography atlas for virtual in vivo dissections". In: *Cortex* 44.8 (Sept. 2008), pp. 1105–1132.

[34] Marco Catani, Derek K. Jones, and Dominic H. ffytche. "Perisylvian language networks of the human brain". In: *Annals of Neurology* 57.1 (2004), pp. 8–16.

[35] Marco Catani et al. "Short frontal lobe connections of the human brain". In: *Cortex* 48.2 (Feb. 2012), pp. 273–291.

[36] Francesco Cauteruccio et al. "Improving QuickBundles to Extract Anatomically Coherent White Matter Fiber-Bundles". In: *Lecture Notes in Computer Science*. Springer Nature, 2016, pp. 633–641.

[37] Eduardo Caverzasi et al. "Identifying preoperative language tracts and predicting postoperative functional recovery using HARDI q-ball fiber tractography in patients with gliomas". In: *Journal of Neurosurgery* 125.1 (July 2016), pp. 33–45.

[38] Eduardo Caverzasi et al. "Q-Ball of Inferior Fronto-Occipital Fasciculus and Beyond". In: *PLoS ONE* 9.6 (June 2014). Ed. by Gaolang Gong, e100274.

[39] Maxime Chamberland et al. "Real-time multi-peak tractography for instantaneous connectivity display". In: *Frontiers in Neuroinformatics* 8 (May 2014).

[40] E. L. Chang et al. “Evaluation of peritumoral edema in the delineation of radiotherapy clinical target volumes for glioblastoma”. In: *Int. J. Radiat. Oncol. Biol. Phys.* 68.1 (May 2007), pp. 144–150.

[41] David Qixiang Chen et al. “Correcting diffusion weight image distortions using anisotropy power maps, a comparative study”. In: *Annual Meeting OHBM*. conference abstract. Organization for Human Brain Mapping. June 2015.

[42] Xiaolei Chen et al. “Prediction of visual field deficits by diffusion tensor imaging in temporal lobe epilepsy surgery”. In: *NeuroImage* 45.2 (Apr. 2009), pp. 286–297.

[43] Olga Ciccarelli et al. “Diffusion-based tractography in neurological disorders: concepts applications, and future developments”. In: *The Lancet Neurology* 7.8 (Aug. 2008), pp. 715–727.

[44] Marc-Alexandre Cote et al. “Cleaning up the mess: tractography outlier removal using hierarchical QuickBundles clustering”. In: *Proceedings of: International Society of Magnetic Resonance in Medicine (ISMRM)* (2015).

[45] S. Crettenand et al. “Quantitative analysis and comparison of diffusion tensor imaging tractography algorithms”. In: *IET Irish Signals and Systems Conference (ISSC 2006)*. Institution of Engineering and Technology (IET), 2006.

[46] I. Cristofori et al. “White and gray matter contributions to executive function recovery after traumatic brain injury”. In: *Neurology* 84.14 (Mar. 2015), pp. 1394–1401.

[47] Andrea L. Crowell et al. “Toward an Understanding of the Neural Circuitry of Major Depressive Disorder Through the Clinical Response to Deep Brain Stimulation of Different Anatomical Targets”. In: *Current Behavioral Neuroscience Reports* 1.2 (Mar. 2014), pp. 55–63.

[48] Michael Dayan, Sylvia Kreutzer, and Chris A. Clark. “Tractography of the optic radiation: a repeatability and reproducibility study”. In: *NMR in Biomedicine* 28.4 (Feb. 2015), pp. 423–431.

[49] Flavio Dell’Acqua et al. “Anisotropic Power Maps: A diffusion contrast to reveal low anisotropy tissues from HARDI data”. In: Session 57 (May 2014). conference abstract. URL: `http://www.ismrm.org/14/program_files/Session57.htm`.

[50] Maxime Descoteaux et al. “Regularized fast, and robust analytical Q-ball imaging”. In: *Magnetic Resonance in Medicine* 58.3 (2007), pp. 497–510.

[51] M. Descoteaux et al. “Deterministic and Probabilistic Tractography Based on Complex Fibre Orientation Distributions”. In: *IEEE Transactions on Medical Imaging* 28.2 (Feb. 2009), pp. 269–286.

[52] Prateek Dixit and G. R. Liu. “A Review on Recent Development of Finite Element Models for Head Injury Simulations”. In: *Archives of Computational Methods in Engineering* (Oct. 2016).

[53] Hugues Duffau. “Interactions Between Diffuse Low-Grade Glioma (DLGG) and Brain Plasticity”. In: *Diffuse Low-Grade Gliomas in Adults*. Springer Nature, 2013, pp. 337–356.

[54] Hugues Duffau. “Stimulation mapping of white matter tracts to study brain functional connectivity”. In: *Nature Reviews Neurology* 11.5 (Apr. 2015), pp. 255–265.

[55] Hugues Duffau. “The anatomo-functional connectivity of language revisited”. In: *Neuropsychologia* 46.4 (2008), pp. 927–934.

[56] Hugues Duffau. “The Dangers of Magnetic Resonance Imaging Diffusion Tensor Tractography in Brain Surgery”. In: *World Neurosurgery* 81.1 (Jan. 2014), pp. 56–58.

[57] Hugues Duffau. “The huge plastic potential of adult brain and the role of connectomics: New insights provided by serial mappings in glioma surgery”. In: *Cortex* 58 (Sept. 2014), pp. 325–337.

[58] A. S. Eden et al. “Emotion Regulation and Trait Anxiety Are Predicted by the Microstructure of Fibers between Amygdala and Prefrontal Cortex”. In: *Journal of Neuroscience* 35.15 (Apr. 2015), pp. 6020–6027.

[59] Sarah M. Fayad et al. “Six-Nine Year Follow-Up of Deep Brain Stimulation for Obsessive-Compulsive Disorder”. In: *PLOS ONE* 11.12 (Dec. 2016). Ed. by Krystof Bankiewicz, e0167875.

[60] Guenther C. Feigl et al. “Magnetic Resonance Imaging Diffusion Tensor Tractography: Evaluation of Anatomic Accuracy of Different Fiber Tracking Software Packages”. In: *World Neurosurgery* 81.1 (Jan. 2014), pp. 144–150.

[61] Juan C. Fernández-Miranda et al. “Asymmetry connectivity, and segmentation of the arcuate fascicle in the human brain”. In: *Brain Structure and Function* 220.3 (Mar. 2014), pp. 1665–1680.

[62] Aaron S. Field et al. “Diffusion tensor eigenvector directional color imaging patterns in the evaluation of cerebral white matter tracts altered by tumor”. In: *Journal of Magnetic Resonance Imaging* 20.4 (2004), pp. 555–562.

[63] Julie A. Fiez, Hanna Damasio, and Thomas J. Grabowski. “Lesion segmentation and manual warping to a reference brain: Intra- and interobserver reliability”. In: *Human Brain Mapping* 9.4 (Apr. 2000), pp. 192–211.

[64] Pierre Fillard et al. “Quantitative evaluation of 10 tractography algorithms on a realistic diffusion MR phantom”. In: *NeuroImage* 56.1 (May 2011), pp. 220–234.

[65] Revanth Reddy Garlapati et al. “More accurate neuronavigation data provided by biomechanical modeling instead of rigid registration”. In: *Journal of Neurosurgery* 120.6 (June 2014), pp. 1477–1483.

[66] Eleftherios Garyfallidis et al. “Dipy a library for the analysis of diffusion MRI data”. In: *Frontiers in Neuroinformatics* 8 (Feb. 2014).

[67] Eleftherios Garyfallidis et al. “QuickBundles a Method for Tractography Simplification”. In: *Frontiers in Neuroscience* 6 (2012).

[68] Eleftherios Garyfallidis et al. “Robust and efficient linear registration of white-matter fascicles in the space of streamlines”. In: *NeuroImage* 117 (Aug. 2015), pp. 124–140.

[69] Helen M. Genova et al. “Facial affect recognition linked to damage in specific white matter tracts in traumatic brain injury”. In: *Social Neuroscience* 10.1 (Sept. 2014), pp. 27–34.

[70] Norman Geschwind. “Disconnexion Syndromes in Animals and Man”. In: *Brain* 88.2 (1965), pp. 237–294.

[71] Norman Geschwind. "Disconnexion Syndromes in Animals and Man". In: *Brain* 88.3 (1965), pp. 585–644.

[72] A. Giese and M. Westphal. "Glioma invasion in the central nervous system". In: *Neurosurgery* 39.2 (Aug. 1996), pp. 235–250.

[73] A. Giese et al. "Cost of Migration: Invasion of Malignant Gliomas and Implications for Treatment". In: *Journal of Clinical Oncology* 21.8 (Apr. 2003), pp. 1624–1636.

[74] Orit A. Glenn et al. "DTI-based three-dimensional tractography detects differences in the pyramidal tracts of infants and children with congenital hemiparesis". In: *Journal of Magnetic Resonance Imaging* 18.6 (Nov. 2003), pp. 641–648.

[75] Alexandra J Golby et al. "Interactive Diffusion Tensor Tractography Visualization for Neurosurgical Planning". In: *Neurosurgery* 68.2 (Feb. 2011), pp. 496–505.

[76] Nelly Gordillo, Eduard Montseny, and Pilar Sobrevilla. "State of the art survey on MRI brain tumor segmentation". In: *Magnetic Resonance Imaging* 31.8 (Oct. 2013), pp. 1426–1438.

[77] Krzysztof Gorgolewski et al. "Nipype: A Flexible Lightweight and Extensible Neuroimaging Data Processing Framework in Python". In: *Frontiers in Neuroinformatics* 5 (2011).

[78] Liberty S. Hamilton et al. "Semi-automated anatomical labeling and inter-subject warping of high-density intracranial recording electrodes in electrocorticography". In: (Submitted 2017).

[79] Janice Hau et al. "Cortical Terminations of the Inferior Fronto-Occipital and Uncinate Fasciculi: Anatomical Stem-Based Virtual Dissection". In: *Frontiers in Neuroanatomy* 10 (May 2016).

[80] Biyu J. He et al. "Breakdown of Functional Connectivity in Frontoparietal Networks Underlies Behavioral Deficits in Spatial Neglect". In: *Neuron* 53.6 (Mar. 2007), pp. 905–918.

[81] R. J. Von Der Heide et al. "Dissecting the uncinate fasciculus: disorders controversies and a hypothesis". In: *Brain* 136.6 (May 2013), pp. 1692–1707.

[82] Roland G Henry et al. “Subcortical pathways serving cortical language sites: initial experience with diffusion tensor imaging fiber tracking combined with intraoperative language mapping”. In: *NeuroImage* 21.2 (Feb. 2004), pp. 616–622.

[83] Guillaume Herbet et al. “Mapping neuroplastic potential in brain-damaged patients”. In: *Brain* 139.3 (Feb. 2016), pp. 829–844.

[84] C. F. Hess et al. “Malignant glioma: patterns of failure following individually tailored limited volume irradiation”. In: *Radiother Oncol* 30.2 (Feb. 1994), pp. 146–149.

[85] Christopher P. Hess et al. “Q-ball reconstruction of multimodal fiber orientations using the spherical harmonic basis”. In: *Magnetic Resonance in Medicine* 56.1 (2006), pp. 104–117.

[86] F. H. Hochberg and A. Pruitt. “Assumptions in the radiotherapy of glioblastoma”. In: *Neurology* 30.9 (Sept. 1980), pp. 907–911.

[87] Hofer. “Reconstruction and dissection of the entire human visual pathway using diffusion tensor MRI”. In: *Frontiers in Neuroanatomy* (2010).

[88] J. Hornak. “Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices”. In: *Brain* 126.7 (Apr. 2003), pp. 1691–1712.

[89] Tamara Ius et al. “Evidence for potentials and limitations of brain plasticity using an atlas of functional resectability of WHO grade II gliomas: Towards a \minimal common brain””. In: *NeuroImage* 56.3 (June 2011), pp. 992–1000.

[90] Sung Ho Jang. “Diffusion Tensor Imaging Studies on Arcuate Fasciculus in Stroke Patients: A Review”. In: *Frontiers in Human Neuroscience* 7 (2013).

[91] Mark Jenkinson et al. “FSL”. In: *NeuroImage* 62.2 (Aug. 2012), pp. 782–790.

[92] Jens H. Jensen et al. “Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging”. In: *Magnetic Resonance in Medicine* 53.6 (2005), pp. 1432–1440.

[93] Ben Jeurissen et al. “Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data”. In: *NeuroImage* 103 (Dec. 2014), pp. 411–426.

[94] Ben Jeurissen et al. “Probabilistic fiber tracking using the residual bootstrap with constrained spherical deconvolution”. In: *Human Brain Mapping* 32.3 (Feb. 2011), pp. 461–479.

[95] Peter Jezzard and Robert S. Balaban. “Correction for geometric distortion in echo planar images from B0 field variations”. In: *Magnetic Resonance in Medicine* 34.1 (July 1995), pp. 65–73.

[96] Fiacro Jiménez et al. “A Patient with a Resistant Major Depression Disorder Treated with Deep Brain Stimulation in the Inferior Thalamic Peduncle”. In: *Neurosurgery* 57.3 (Sept. 2005), pp. 585–593.

[97] Derek K. Jones, Thomas R. Knosche, and Robert Turner. “White matter integrity fiber count, and other fallacies: The do’s and don’ts of diffusion MRI”. In: *NeuroImage* 73 (June 2013), pp. 239–254.

[98] Kesshi Jordan. *Publication Repository*. `https://github.com/kesshijordan/Publication_Repository/cci.py`. 2017.

[99] Kesshi Jordan et al. “Cluster Confidence Index: A Streamline-wise Pathway Reproducibility Metric for Diffusion-Weighted MRI Tractography”. In: *Journal of Neuroimaging* (Accepted 2017).

[100] P. Kalanithi and A. Verghese. *When breath becomes air*. 2016.

[101] Anisha Keshavan et al. “Mindcontrol: A Web Application for Brain Segmentation Quality Control”. In: *NeuroImage* (Mar. 2017).

[102] S. H. Kim and S. H. Jang. “Prediction of Aphasia Outcome Using Diffusion Tensor Tractography for Arcuate Fasciculus in Stroke”. In: *American Journal of Neuroradiology* 34.4 (Oct. 2012), pp. 785–790.

[103] Manabu Kinoshita et al. “Fiber-tracking does not accurately estimate size of fiber bundle in pathological condition: initial neurosurgical experience using neuronavigation and subcortical white matter stimulation”. In: *NeuroImage* 25.2 (Apr. 2005), pp. 424–429.

[104] Masashi Kinoshita et al. “Chronic spatial working memory deficit associated with the superior longitudinal fasciculus: a study using voxel-based lesion-symptom mapping and intraoperative direct stimulation in right prefrontal glioma surgery”. In: *Journal of Neurosurgery* 125.4 (Oct. 2016), pp. 1024–1032.

[105] Pradeep Kumar et al. “Prediction of upper extremity motor recovery after subacute intracerebral hemorrhage through diffusion tensor imaging: a systematic review and meta-analysis”. In: *Neuroradiology* 58.10 (2016), pp. 1043–1050.

[106] Pradeep Kumar et al. “Prediction of upper limb motor recovery after subacute ischemic stroke using diffusion tensor imaging: a systematic review and meta-analysis”. In: *Journal of stroke* 18.1 (2016), p. 50.

[107] Michel Lacroix et al. “A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival”. In: *Journal of Neurosurgery* 95.2 (2001), pp. 190–198.

[108] Denis Le Bihan et al. “Diffusion tensor imaging: concepts and applications”. In: *Journal of magnetic resonance imaging* 13.4 (2001), pp. 534–546.

[109] Delphine Leclercq et al. “Comparison of diffusion tensor imaging tractography of language tracts and intraoperative subcortical stimulations”. In: *Journal of Neurosurgery* 112.3 (Mar. 2010), pp. 503–511.

[110] Alexander Leemans and Derek K. Jones. “The B-matrix must be rotated when correcting for subject motion in DTI data”. In: *Magnetic Resonance in Medicine* 61.6 (June 2009), pp. 1336–1349.

[111] Stéphane Lehéricy et al. “Presurgical Functional Localization Possibilities Limitations, and Validity”. In: *Clinical Functional MRI*. Springer Nature, Jan. 2014, pp. 247–267.

[112] Stéphane Lehéricy et al. “Validity of Presurgical Functional Localization”. In: *Clinical Functional MRI*. Springer Nature, 2007, pp. 167–187.

[113] J. F. Lehman et al. “Rules Ventral Prefrontal Cortical Axons Use to Reach Their Targets: Implications for Diffusion Tensor Imaging Tractography and Deep Brain Stimulation for Psychiatric Illness”. In: *Journal of Neuroscience* 31.28 (July 2011), pp. 10392–10402.

[114] Andres M. Lozano et al. “Subcallosal Cingulate Gyrus Deep Brain Stimulation for Treatment-Resistant Depression”. In: *Biological Psychiatry* 64.6 (Sept. 2008), pp. 461–467.

[115] J. Luis Lujan et al. “Tractography-Activation Models Applied to Subcallosal Cingulate Deep Brain Stimulation”. In: *Brain Stimulation* 6.5 (Sept. 2013), pp. 737–739.

[116] N. Makris et al. “Delineation of the Middle Longitudinal Fascicle in Humans: A Quantitative In Vivo, DT-MRI Study”. In: *Cerebral Cortex* 19.4 (July 2008), pp. 777–785.

[117] N. Makris et al. “Human middle longitudinal fascicle: variations in patterns of anatomical connections”. In: *Brain Structure and Function* 218.4 (July 2012), pp. 951–968.

[118] M. L. Mandelli et al. “Frontal White Matter Tracts Sustaining Speech Production in Primary Progressive Aphasia”. In: *Journal of Neuroscience* 34.29 (July 2014), pp. 9754–9767.

[119] Maria Luisa Mandelli et al. “Quantifying accuracy and precision of diffusion MR tractography of the corticospinal tract in brain tumors”. In: *Journal of Neurosurgery* 121.2 (Aug. 2014), pp. 349–358.

[120] Daniel Martinez-Ramirez et al. “Update on deep brain stimulation in Parkinson’s disease”. In: *Translational Neurodegeneration* 4.1 (June 2015).

[121] Juan Martino et al. “Anatomic dissection of the inferior fronto-occipital fasciculus revisited in the lights of brain stimulation data”. In: *Cortex* 46.5 (2010), pp. 691–699.

[122] Ioannis N. Mavridis. “Commentary: Tractography-Activation Models Applied to Subcallosal Cingulate Deep Brain Stimulation”. In: *Frontiers in Neuroanatomy* 9 (Nov. 2015).

[123] Bjoern H. Menze et al. “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)”. In: *IEEE Transactions on Medical Imaging* 34.10 (Oct. 2015), pp. 1993–2024.

[124] Sarah Meyer et al. “Voxel-based lesion-symptom mapping of stroke lesions underlying somatosensory deficits”. In: *NeuroImage: Clinical* 10 (2016), pp. 257–266.

[125] Susumu Mori and Peter van Zijl. “Fiber tracking: principles and strategies–a technical review”. In: *NMR in Biomedicine* 15.7-8 (2002), pp. 468–480.

[126] Peter F. Neher et al. “Strengths and weaknesses of state of the art fiber tractography pipelines – A comprehensive in-vivo and phantom evaluation study using Tractometer”. In: *Medical Image Analysis* 26.1 (Dec. 2015), pp. 287–305.

[127] Christopher Nimsky et al. “Preoperative and Intraoperative Diffusion Tensor Imaging-based Fiber Tracking in Glioma Surgery”. In: *Neurosurgery* 56.1 (Jan. 2005), pp. 130–138.

[128] Lauren J. O’Donnell et al. “Automated white matter fiber tract identification in patients with brain tumors”. In: *NeuroImage: Clinical* 13 (2017), pp. 138–153.

[129] Michael S Okun et al. “Subthalamic deep brain stimulation with a constant-current device in Parkinson’s disease: an open-label randomised controlled trial”. In: *The Lancet Neurology* 11.2 (Feb. 2012), pp. 140–149.

[130] Sinisa Pajevic and Carlo Pierpaoli. “Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: Application to white matter fiber tract mapping in the human brain”. In: *Magnetic Resonance in Medicine* 42.3 (Sept. 1999), pp. 526–540.

[131] Costanza Papagno et al. “Connectivity constraints on cortical reorganization of neural circuits involved in object naming”. In: *NeuroImage* 55.3 (Apr. 2011), pp. 1306–1313.

[132] Franco Pestilli et al. “Evaluation and statistical inference for human connectomes”. In: *Nature Methods* 11.10 (Sept. 2014), pp. 1058–1063.

[133] Carlo Pierpaoli and Peter J. Basser. “Toward a quantitative assessment of diffusion anisotropy”. In: *Magnetic Resonance in Medicine* 36.6 (Dec. 1996), pp. 893–906.

[134] J. L. PRICE. "Definition of the Orbital Cortex in Relation to Specific Connections with Limbic and Visceral Structures and Other Cortical Regions". In: *Annals of the New York Academy of Sciences* 1121.1 (Sept. 2007), pp. 54–71.

[135] Patricio Riva-Posse et al. "Defining Critical White Matter Pathways Mediating Successful Subcallosal Cingulate Deep Brain Stimulation for Treatment-Resistant Depression". In: *Biological Psychiatry* 76.12 (Dec. 2014), pp. 963–969.

[136] P Riva-Posse et al. "A connectomic approach for subcallosal cingulate deep brain stimulation surgery: prospective targeting in treatment-resistant depression". In: *Molecular Psychiatry* (Apr. 2017).

[137] P. Riva-Posse et al. "Target selection in subcallosal cingulate deep brain stimulation: a prospective connectomic-based method". In: *Brain Stimulation* 10.2 (Mar. 2017), p. 482.

[138] Ariel Rokem et al. "Evaluating the Accuracy of Diffusion MRI Models in White Matter". In: *PLOS ONE* 10.4 (Apr. 2015). Ed. by Yun Wang, e0123272.

[139] Edmund T. Rolls. "The roles of the orbitofrontal cortex via the habenula in nonreward and depression and in the responses of serotonin and dopamine neurons". In: *Neuroscience & Biobehavioral Reviews* 75 (Apr. 2017), pp. 331–334.

[140] V. Salanova et al. "Long-term efficacy and safety of thalamic stimulation for drug-resistant partial epilepsy". In: *Neurology* 84.10 (Feb. 2015), pp. 1017–1025.

[141] Ana Sanjuán et al. "Automated identification of brain tumors from single MR images based on segmentation with refined patient-specific priors". In: *Frontiers in Neuroscience* 7 (2013).

[142] Alexander Sartorius and Fritz A. Henn. "Deep brain stimulation of the lateral habenula in treatment resistant major depression". In: *Medical Hypotheses* 69.6 (Jan. 2007), pp. 1305–1308.

[143] Thomas E. Schlaepfer. "Deep Brain Stimulation for Major Depression—Steps on a Long and Winding Road". In: *Biological Psychiatry* 78.4 (Aug. 2015), pp. 218–219.

[144] Thomas E Schlaepfer et al. "Deep Brain Stimulation of the Human Reward System for Major Depression—Rationale Outcomes and Outlook". In: *Neuropsychopharmacology* 39.6 (Feb. 2014), pp. 1303–1314.

[145] Michel Thiebaut de Schotten et al. "Monkey to human comparative anatomy of the frontal lobe association tracts". In: *Cortex* 48.1 (Jan. 2012), pp. 82–96.

[146] Kristin K. Sellers et al. "Electrical Stimulation of Human Orbitofrontal Cortex Engages Limbic Circuits and Acutely Improves Negative Mood". In: (Submitted 2017).

[147] Rebecca L. Siegel, Kimberly D. Miller, and Ahmedin Jemal. "Cancer statistics, 2017". In: *CA: A Cancer Journal for Clinicians* 67.1 (2017), pp. 7–30. ISSN: 1542-4863.

[148] S Smith and T Nichols. "Threshold-free cluster enhancement: Addressing problems of smoothing threshold dependence and localisation in cluster inference". In: *NeuroImage* 44.1 (Jan. 2009), pp. 83–98.

[149] Stephen M. Smith. "Fast robust automated brain extraction". In: *Human Brain Mapping* 17.3 (Nov. 2002), pp. 143–155.

[150] Stephen M. Smith et al. "Advances in functional and structural MR image analysis and implementation as FSL". In: *NeuroImage* 23 (Jan. 2004), S208–S219.

[151] Roger Stupp et al. "Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma". In: *New England Journal of Medicine* 352.10 (2005), pp. 987–996.

[152] J-Donald Tournier, Fernando Calamante, and Alan Connelly. "MRtrix: Diffusion tractography in crossing fiber regions". In: *International Journal of Imaging Systems and Technology* 22.1 (Feb. 2012), pp. 53–66.

[153] J-Donald Tournier, Fernando Calamante, and Alan Connelly. "Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution". In: *NeuroImage* 35.4 (May 2007), pp. 1459–1472.

[154] J.-Donald Tournier et al. "Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution". In: *NeuroImage* 23.3 (Nov. 2004), pp. 1176–1185.

[155] Antonio Tristán-Vega and Santiago Aja-Fernández. "DWI filtering using joint information for DTI and HARDI". In: *Medical Image Analysis* 14.2 (Apr. 2010), pp. 205–218.

[156] Antonio Tristán-Vega and Carl-Fredrik Westin. "Probabilistic ODF Estimation from Reduced HARDI Data with Sparse Regularization". In: *Lecture Notes in Computer Science*. Springer Nature, 2011, pp. 182–190.

[157] Antonio Tristán-Vega, Carl-Fredrik Westin, and Santiago Aja-Fernández. "Estimation of fiber Orientation Probability Density Functions in High Angular Resolution Diffusion Imaging". In: *NeuroImage* 47.2 (Aug. 2009), pp. 638–650.

[158] Kwong Ming Tse et al. "Development and validation of two subject-specific finite element models of human head against three cadaveric experiments". In: *International Journal for Numerical Methods in Biomedical Engineering* 30.3 (Nov. 2013), pp. 397–415.

[159] David S. Tuch. "Q-ball imaging". In: *Magnetic Resonance in Medicine* 52.6 (2004), pp. 1358–1372.

[160] David S. Tuch et al. "Diffusion MRI of Complex Neural Architecture". In: *Neuron* 40.5 (Dec. 2003), pp. 885–895.

[161] David S. Tuch et al. "High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity". In: *Magnetic Resonance in Medicine* 48.4 (Sept. 2002), pp. 577–582.

[162] Setsu Wakana et al. "Reproducibility of quantitative tractography methods applied to cerebral white matter". In: *NeuroImage* 36.3 (July 2007), pp. 630–644.

[163] Ruopeng Wang and Van J. Wedeen. *TrackVis*. 2015. URL: `http://www.trackvis.org/`.

[164] R Wang et al. "Diffusion toolkit: a software package for diffusion imaging data processing and tractography". In: *Proc Intl Soc Mag Reson Med*. Vol. 15. 3720. 2007.

[165] Rene Werner et al. “Beyond cost function masking: RPCA-based non-linear registration in the context of VLSM”. In: *2016 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*. Institute of Electrical and Electronics Engineers (IEEE), June 2016.

[166] Mark W. Woolrich et al. “Bayesian analysis of neuroimaging data in FSL”. In: *NeuroImage* 45.1 (Mar. 2009), S173–S186.

[167] Jin-Song Wu et al. “Clinical Evaluation and Follow-Up Outcome of Diffusion Tensor Imaging-Based Functional Neuronavigation”. In: *Neurosurgery* 61.5 (Nov. 2007), pp. 935–949.

[168] Kei Yamada et al. “MR Tractography: A Review of Its Clinical Applications”. In: *Magnetic Resonance in Medical Sciences* 8.4 (2009), pp. 165–174.

[169] Jason D. Yeatman et al. “Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification”. In: *PLoS ONE* 7.11 (Nov. 2012). Ed. by Christian Beaulieu, e49790.

[170] Edward H. Yeterian et al. “The cortical connectivity of the prefrontal cortex in the monkey brain”. In: *Cortex* 48.1 (Jan. 2012), pp. 58–81.

[171] M. Yogarajah et al. “Defining Meyer’s loop-temporal lobe resections visual field deficits and diffusion tensor tractography”. In: *Brain* 132.6 (May 2009), pp. 1656–1668.

[172] Sang Wook Yoo et al. “An Example-Based Multi-Atlas Approach to Automatic Labeling of White Matter Tracts”. In: *PLOS ONE* 10.7 (July 2015). Ed. by Alessandro Gozzi, e0133337.

[173] H Zhang et al. “Deformable registration of diffusion tensor MR images with explicit orientation optimization”. In: *Medical Image Analysis* 10.5 (Oct. 2006), pp. 764–785.

## Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature Kesshi M. Jordan Date June 12, 2017