HAL
open science

The influence of regions of interest on tractography virtual dissection protocols: general principles to learn and to follow

Francois Rheault, Kurt G Schilling, Sami Obaid, John Begnoche, Laurie E Cutting, Maxime Descoteaux, Bennett A Landman, Laurent Petit

To cite this version:

Francois Rheault, Kurt G Schilling, Sami Obaid, John Begnoche, Laurie E Cutting, et al.. The influence of regions of interest on tractography virtual dissection protocols: general principles to learn and to follow. Brain Structure and Function, 2022, 227 (6), pp.2191-2207. 10.1007/s00429-022-02518-6 . hal-03781913

HAL Id: hal-03781913

https://hal.science/hal-03781913v1

Submitted on 20 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Manuscript

Click here to access/download;Manuscript;tractostorm_v2_ROIs_main_manus

Click here to view linked References

The influence of regions of interest on tractography virtual dissection protocols: general principles to learn and to follow

Francois Rheault[*,a], Kurt G. Schilling[b,c], Sami Obaid[d,e], John Begnoche[f], Laurie E. Cutting[g], Maxime Descoteaux[d], Bennett A. Landman[a,b,c,h] Laurent Petit[i]

[a]*Electrical and Computer Engineering, Vanderbilt University, Nashville, USA;*
[b]*Vanderbilt University Institute of Imaging, Nashville, USA;*
[c]*Department of Radiology and Radiological Science, Vanderbilt University Medical Center, Nashville, USA*
[d]*Sherbrooke Connectivity Imaging Laboratory (SCIL), Département d'Informatique, Université de Sherbrooke, Canada;*
[e]*University of Montreal, Health Center Research Center, Montreal, Canada;*
[f]*The Center for Cognitive Medicine, Department of Psychiatry, Vanderbilt University Medical Center, Nashville, USA;*
[g]*Vanderbilt Kennedy Center, Vanderbilt, University Medical Center, Nashville, USA;*
[h]*Computer Science, Vanderbilt University, Nashville, USA;*
[i]*Groupe d'Imagerie Neurofonctionnelle, Institut Des Maladies Neurodégénératives, CNRS, CEA University of Bordeaux, Bordeaux, France*

Abstract

Efficient communication across fields of research is challenging, especially when they are at opposite ends of the physical and digital spectrum. Neuroanatomy and neuroimaging may seem close to each other, but the terminology and processes to study the brain can be very different. More specifically, investigations of white matter anatomy are susceptible to this challenge. This gap creates disagreement on ways to define the same underlying anatomy. Even when trying to isolate the same structure, according to a specific anatomical definition, it is a non-trivial task to convert the neuroanatomical knowledge to instructions and rules to be executed in neuroimaging software. In the process called "virtual dissection" used to isolate coherent white matter structure in tractography, each white matter pathway has its own set of landmarks (regions of interest) used as inclusion and exclusion criteria. The ability to reproducibly segment and study these pathways is critical for scientific progress, yet, variability may depend on region placement, and the investigator placing the region (i.e a rater).When rater variability is taken into account, the impact made by each region of interest becomes even more difficult to interpret. A delicate balance between anatomical validity, impact on the virtual dissection and raters reproducibility emerge. In this work, we investigate this balance by leveraging manual delineation data of a group of raters from a previous study to quantify which set of landmarks and criteria contribute most to variability in virtual dissection. To supplement our analysis, the variability of each pathway with a region-by-region exploration was performed. We present a detailed exploration and description of each region, the causes of variability, its impacts and potential solutions for future protocols. Finally, we provide a brief overview of the lessons learned from our previous virtual dissection projects and propose recommendations for future virtual dissection protocols as well as perspectives to reach better community agreement when it comes to anatomical definitions of white matter pathways.

1. Introduction

Describing the complex architecture of the white matter of the human brain is a challenging task. There is a balance between neuroanatomical accuracy, simplicity and conciseness when one attempts to communicate the knowledge about the cortical terminations, the landmarks surrounding pathways or their shapes (Benedictis et al., 2016; David et al., 2019; Dick and Tremblay, 2012; Fernández-Miranda et al., 2015; Hau et al., 2017; Maffei et al., 2018; Vavassori et al., 2021). Diffusion MRI tractography facilitates the process of designing more efficient ways to convey neuroanatomical descriptions at a large scale. Virtual reconstruction of major WM pathways with tractography is a useful tool to probe structural connectivity and has the potential to be useful for teaching neuroanatomy.

Virtual dissection allows the isolation of pathways of interest by targeting streamlines using inclusion and exclusion regions of interest (ROIs). The instructions to delineate a specific pathway can be detailed and shared using a task protocol. However, tractography reconstructions do not always accurately represent the underlying anatomy (Jbabdi and Johansen-Berg, 2011; Maier-Hein et al., 2017; O'Donnell and Pasternak, 2015; Rheault et al., 2020b; Schilling et al., 2020). It is important to consider that the algorithmic limitations of tractography will impact virtual dissections in various ways. If inclusion ROIs are too large or exclusion ROIs too small, bundles can be riddled with spurious streamlines (artificially inflating the volume). If inclusion ROIs are too small or exclusion ROIs too large, bundles can appear sparse or disappear entirely (giving the appearance of a damaged/obliterated pathway). If the virtual dissection is not anatomically informed or if the ROIs are not well-positioned, the observed bundle can be a false positive in its entirety.

These challenges to virtual dissection can be addressed by creating protocols that detail the segmentation tasks and target specific situations and audiences. For example, a protocol designed for neurosurgical exploration executed by experts in neuroanatomy can be very different from an application in aging with thousands of subjects performed by data scientists. There is nothing that prevents virtual dissection protocols from requiring preprocessing (e.g. Freesurfer (Dale et al., 1999)) to better handle specific situations. However, depending on the intended use or targeted audience it may be ill-advised to rely on external software to provide ROIs for the virtual dissection process. For example, limitations due to pathologies or injuries where the external software fails (e.g. tumors, strokes, etc.) can limit its scope. Limitations due to the technical knowledge or processing capacity of the target audience can also limit the virtual dissection protocol's scope. It is important to explicitly mention that protocols cannot be (easily) generalized to a new interactive segmentation software or a protocol designed for diffusion tensor imaging (DTI) is not necessarily valid when applied to a probabilistic tractogram obtained from constrained spherical deconvolution (CSD) or any other high angular resolution diffusion imaging (HARDI).

Virtual dissection protocols must take into account the scenarios for which they are designed and the previously mentioned biases to limit wrongful results and erroneous interpretations. There is a delicate balance between maintaining anatomical accuracy and avoiding spurious streamlines (Schilling et al., 2020; Thomas et al., 2014). To avoid inaccuracies (i.e anatomically implausible streamlines), virtual dissection protocols must target known pathways with unambiguous instructions. Major challenges of protocol design include: clearly describing ROIs position and shape, including only useful/relevant ROIs with predictable effects and anticipating the "spectrum" of spatial effects on bundles caused by ROIs (and if they all are anatomically plausible).

In this work, we investigate (1) the variability in region delineation, (2) the importance of ROIs and ROIs placement on resulting bundles and finally (3) we visualize the spatial impact caused by ROIs' variability. We made observations that allow us to extract recommendations that could help to achieve "clear instructions" and improve future protocols to make them more robust. With the careful examination of each ROI intrinsic variability and the effect it has on pathways, recommendations to increase anatomical accuracy, reduce variability and decrease workload for raters were derived. Furthermore, the insight obtained from our investigation can be consolidated to design better ways to *teach* neuroanatomy from tractography (Arantes et al., 2018; Chang and Molnár, 2015). It is noteworthy that the goal of this work is **not** to design a perfect fits-all protocol, **not** to claim it can be generalized to pathology and **not** to argue that the anatomical definition is better or more valid than others. The aim of this project is first and foremost to gain insight into white matter pathway segmentation using standardized protocols and how to design, share and evaluate them.

1.1 Background

From a neuroimaging perspective, ex vivo approaches such as Klingler's dissection (Agrawal et al., 2011), histological staining/tracing (Alturkistani et al., 2016; Saleeba et al., 2019) or more recent methods such as polarized light imaging (PLI) or optical coherence tomography (OCT) (Axer et al., 2016; Wang et al., 2015) are considered *ground truth* to reconstruction methods from diffusion MRI tractography. However, they should be seen more like a *gold standard* with several caveats to be acknowledged. Large-scale patterns can be agreed upon, but each laboratory has its techniques. This limits the ability to reach an agreement on the anatomical definition. Even when knowing that two regions are indeed connected, the exact path is not as clearly defined as some would expect. While these challenges are relevant to neuroimaging, they must be resolved by neuroanatomists. However, we can design virtual dissection protocols based on definitions even if they are not considered a consensus and explore how to convey such information as well as how to effectively execute virtual dissection tasks.

The reproducibility of virtual dissection protocols can be evaluated even without a strong consensus on the anatomical definition (Rheault et al., 2020a; Schilling et al., 2021a). This is a time-consuming process that requires collaborations from multiple individuals, the careful design of a virtual dissection protocol and a rigorous framework for the analysis. Rheault et al., 2020

(Tractostorm 1) proposed guidelines for robust quantification of reproducibility. By assembling a group of non-experts and experts (with and without a formal background in anatomy) and providing them with triplicate datasets (5 HCP subjects x 3 copies), variability in the virtual dissection of the pyramidal tracts (Chenot et al., 2019) was quantified. It was shown that experts and non-experts had similar intra-rater and inter-rater (within their respective group) reproducibility scores.

Due to the heavy workload and limitations present in the first project, this approach was expanded with Rheault et al., 2022b (Tractostorm 2). The study design was modified to investigate agreement between raters over time (by using 5 HCP subjects x 4 copies in a specific order), an online course was added to introduce the virtual dissection and the number of bundles targeted by the protocol was increased (to 3). This work revealed that raters' agreement did not significantly change over time, showed a significant (global) increase in reproducibility as a result of the online course and confirmed that each bundle of interest has its own reproducibility scores and trends.

From the work of Rheault et al., 2022b (Tractostorm 2) manual delineation of WM landmarks (15) from 20 raters on 20 datasets (5 HCP subjects x 4 copies) were obtained. This effort led to 6000 regions of interest (ROIs) that were not detailed in the analysis presented in Rheault et al., 2022b. The current work presents this previous set of data by investigating the observed **intrinsic** variability (position, shape) from the manual delineations as well as the impact on each pathway **caused by the regions of interest** involved in the pathway's virtual dissection.

Rheault et al., 2022a first proposed a distinction between a theoretical definition and a practical definition. The term "**Theoretical definition"** was introduced as *"must refer to conceptual landmarks and regions, convey the general shape, orientation and terminations. They must aim to be distinct enough between bundles while encompassing a variety of practical definitions. Such definitions should refer to anatomical structures of the brain and not be tied to specific contrasts or processing"*. **Practical definitions** are analogous to virtual dissection protocols and were introduced as *"to a specific execution of a theoretical definition. They can change according to the target audience (e.g. experts vs. nonexperts), time requirements (e.g. quick approximation vs. careful delineation), visualization software (e.g. MI-Brain,* Rheault et al., 2016 ***vs****. TrackVis,* Wang et al., 2007 ***vs****. DSI-Studio,* Yeh, 2021*), or underlying processing (e.g. DTI* ***vs****. HARDI, deterministic vs. probabilistic). This also applies to automatic dissection methods"*.

The virtual dissection protocol used as a basis for this work is from Rheault et al., 2022b and the current protocol (Tractostorm 2 - Zenodo) is considered a practical definition.

2. Methods

All data used for this work comes from Tractostorm 2 (Rheault et al., 2022b). In summary, the study design involved 20 raters that contributed by carrying out the **same pre-established** protocol. Each rater had 3 bundle dissections to perform for each of the 20 datasets. The 20 datasets were, in fact, 5 HCP subjects each with 4 duplicates each. In total, 1200 bundles and 6000 regions of interest were submitted.

This specific protocol used by all raters in this study (available at Tractostorm 2 - Zenodo) was designed for nonexperts in neuroanatomy (basic knowledge of brain organization) with some level of familiarity with tractography. The protocol is adapted for probabilistic tractography from HARDI local models (Garyfallidis et al., 2014; Tournier et al., 2019) and can be carried out in approximately 30 minutes without external processing apart from the tractography itself. The result of the virtual dissection protocol is a set of 3 bundles of interest (BOI): One commissural (body of the corpus callosum, CC), one association (left arcuate fasciculus, AF) and one projection (left pyramidal tract, PYT). The expected shape of the pathways can be seen in Figure 1, the rendering was generated from the *average* BOIs from the submissions (voxel-wise majority vote).

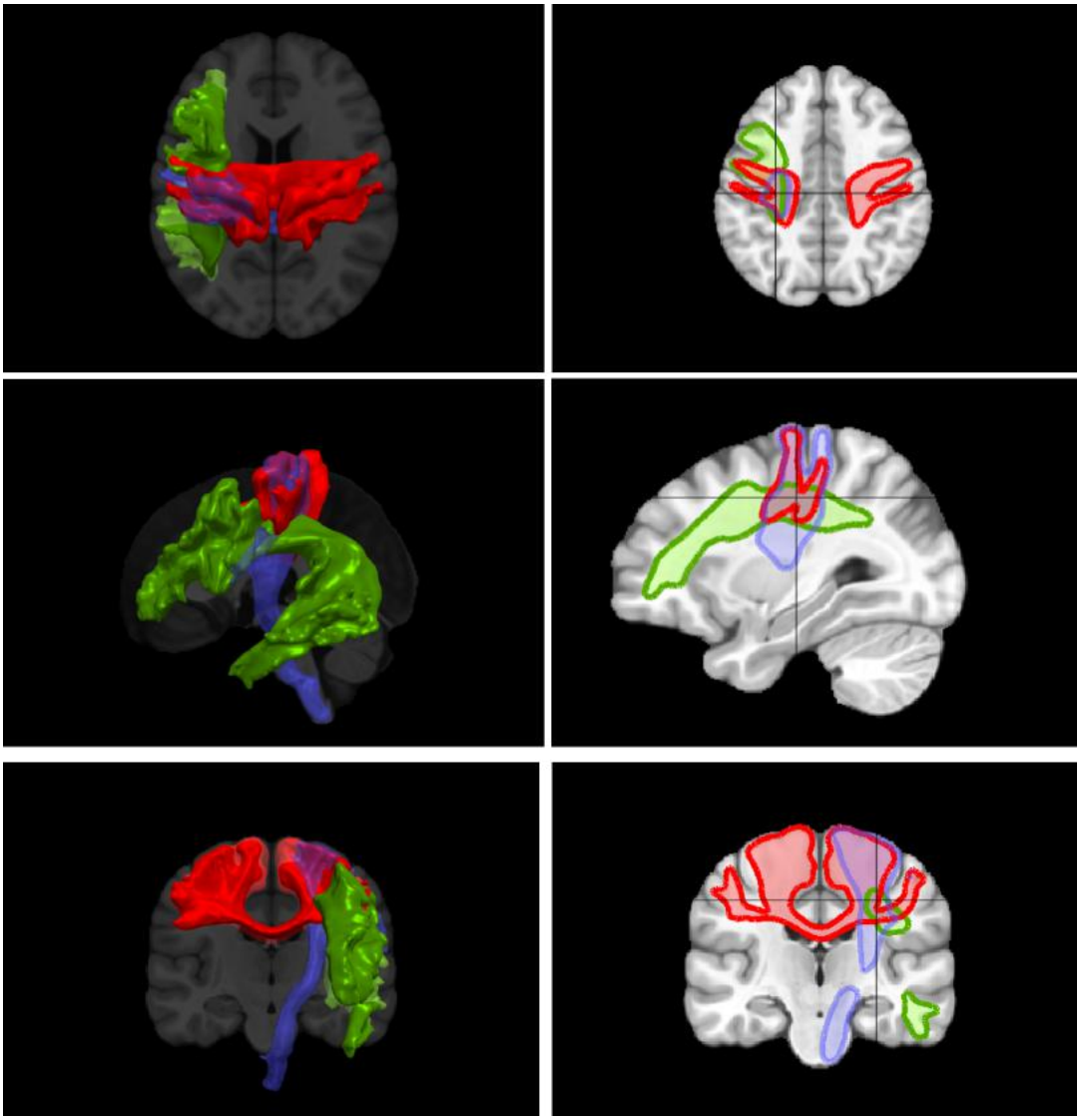

Figure 1: Average expected shape of bundles segmented using the protocol. The corpus callosum, arcuate fasciculus and pyramidal tract in red, green and blue respectively. The choice of bundles was made to explore the 3 main types of pathways (commissural, association and projection) as well as to limit the workload on our collaborators. The isosurfaces were generated at the probability value of 0.5 based on segmentation performed by the principal investigators. The protocol provided to collaborators/raters (available at Tractostorm 2 - Zenodo) presented that figure to provide context to the tasks they were about to initiate.

A total of 15 regions of interest had to be delineated (some are used for more than one pathway) for the project. There were three main categories of ROIs: Disk region, large hand-drawn plane, whole slice plane. The decision to have only planar regions representing simple shapes was made in an attempt to accelerate the virtual dissection process, facilitate interpretation and keep variability to a minimum. Each ROI had to be anatomically justified and was designed to leave as little as possible to interpretation. Figure 2 shows the 15 ROIs from a random rater on the first HCP dataset. The full protocol description is available online.

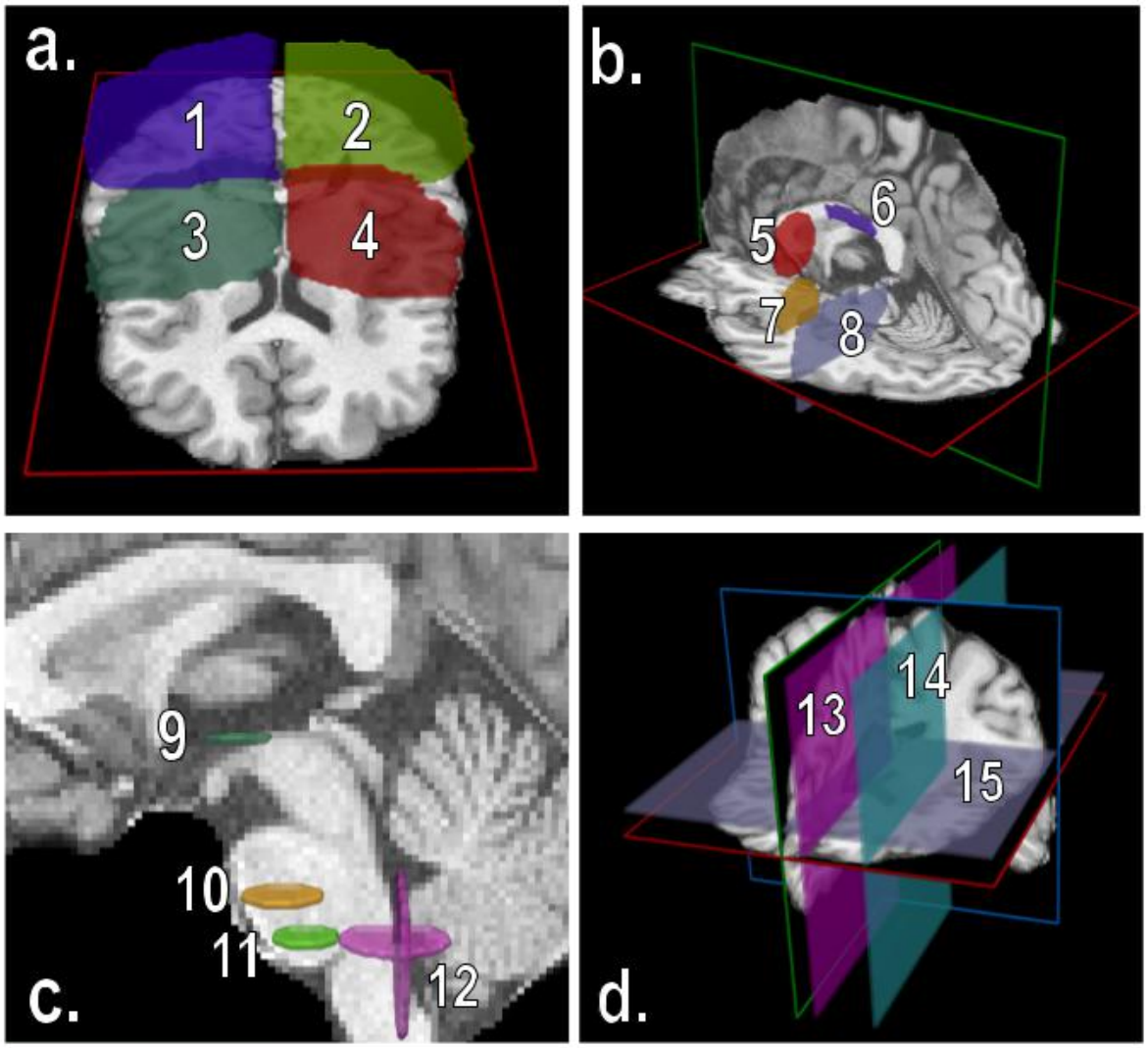

Figure 2: Rendering of the ROIs from a single HCP subject showing their expected size, shape and position. In **a.**, the postcentral (1-2) and precentral (3-4) (left and right), are defined as planes. The planes represent the most anterior region associated with the precentral and the most posterior region associated with the postcentral. These two regions aimed to define the space in between as where the entire precentral and postcentral gyri are located. In **b.**, the centrum semiovale (5), body of the corpus callosum (6), the temporal stem (7) and the temporal lobe entry (8). The regions are defined either as circular-shaped or hand-drawn. In **c.**, the internal capsule (9), midbrain (10), medulla oblongata (11) and cerebellum entry (12). The region's names do not exactly reflect their position, but rather where tractography is expected to reach (or go through) when filtered by the ROI. In **d.**, the mid-sagittal plane (13), the medial sagittal limit (14) and the lower axial limit (15) are 3 planar regions to prevent tractography from reaching a specific area (for more details, Tractostorm 2 - Zenodo). Segmentation of the corpus callosum involved ROI #1-2-3-4-6-15, segmentation of the arcuate fasciculus involved ROI #2-4-5-7-8-14, segmentation of the pyramidal tract involved ROI #2-4-9-10-11-12-13.

2.1 Analysis

To investigate the intrinsic variability of the positioning and shape of the ROIs, we computed the distance between the center of mass as well as a Hausdorff distance (shape similarity, Huttenlocher et al., 1993) for matching ROIs across raters. These metrics relay information about the disagreement in position (or *difficulty* to position the ROIs) and the disagreement in shape. For the circular ROIs, the disagreement in shape is analogous to variation in the diameter chosen by the rater in the software (Rheault et al., 2016).

To evaluate the impact of ROIs variability on the pathways, we used two approaches: Removal of a ROI from a set of dissection criteria and substitution of ROIs between raters.

ROIs used in our virtual dissection protocol are not necessarily all created equal. Some were designed to enforce a known anatomical rule, others to constrain to a specific shape to avoid spurious streamlines. However, ROIs interact with each other, leading to redundant criteria and small contributions to the overall virtual dissection. To quantify the impact of a ROI variability on a pathway the virtual dissection is performed without one of the ROI and compares it to the baseline described above. This operation informs us of the overall **usefulness** of a specific ROI in the virtual dissection protocol.

The substitution approach uses all the ROIs to dissect a BOI to obtain a baseline. Then, the ROI is substituted with one from another rater and is finally compared to the baseline using the bundle adjacency distance (Garyfallidis et al., 2018; Rheault et al., 2020a). This operation informs us on the **influence** a specific ROI's intrinsic variability has on the final result (bundle) and if this influence is the same across the 3 bundles.

The operations are repeated for each ROI, across raters and for each bundle of interest. Since some ROIs are used by more than one bundle, this leads to 19 distributions (CC: 6, AF: 6, PYT: 7) from 15 ROIs.

Finally, an ROI-by-ROI exploration of each BOI's variability caused by each ROI was performed. This material (available in the supplementary materials) provides a detailed description of each region, the causes of variability, its impacts and potential solutions. For each ROI, a figure will accompany the written description. Figure 3 shows the visualization method used to inform how the variability of ROIs translates to the bundles. The color-encoding of streamlines was designed to facilitate visualization of the impact of rater variability; this figure shows how an average bundle is impacted by the application of all ROIs. If a streamline is selected/excluded by 1,2 and up to 80 ROIs this score is mapped to a color spectrum (lower right). Purple means that the streamlines are consistently involved in the segmentation by the raters and yellow means only a few raters selected/excluded the streamlines. The percentage represents the ratio of raters that missed the streamline; 1/80 (purple): almost no rater missed that streamline, 79/80 (yellow): almost all raters missed that streamline).

All summary notes are all organized similarly. The caption of Figure 3 contains an example of interpretations that are drawn from the visualization. The written text aims to describe:

- The importance and meaning of the anatomical ROI
- How the raters were expected to locate and draw the ROI
- What were the sources of confusion and why did it cause so much/little variability
- What are the ways to reduce variability and is the ROI necessary
- What are the consequences if a rater misinterpreted the instructions

This meticulous inspection combined with the previously mentioned analysis are the basis for our summarized recommendation of Section 3.2 and conclusions. Only an overview of one ROI per bundle is presented in the results section. All 19 summary notes are available in the supplementary materials.

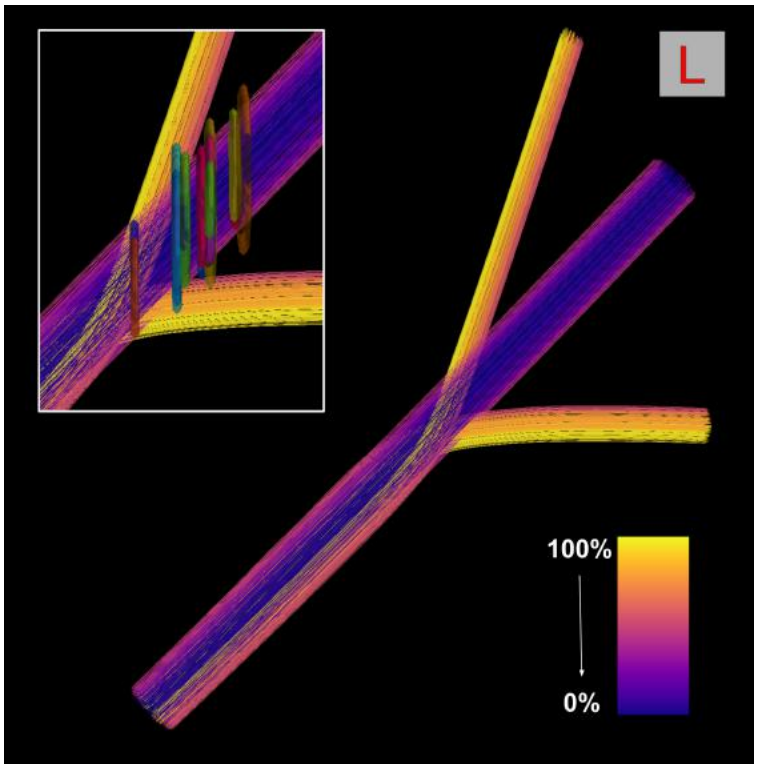

Figure 3: In this figure, we can observe that the (inclusion) ROIs are generally well-positioned in the main axis of the bundle. However, some are placed “too early” and partially intersect the crossing and the effect is that branches are accidentally selected. Variation in shape/size seems to have a minor impact as most of the center bundle is correctly selected. **Typically, the more yellow present in the streamlines, the more impactful (and potentially prone to variability) the ROI is in the set of rules to segment the bundles.** For example, we can conclude that adding a few millimeters after the crossing as a safe margin for the segmentation would increase reproducibility and guarantee that the center of the bundle would be well-defined. The choice to use a subcomponent of the Fibercup (Poupon et al., 2008) is simply to help navigation and illustrate the organization/coloration of streamlines.

2. Results

2.1 Regions of interest intrinsic variability

Variability of the regions of interest positioning is different for each ROIs. Despite a low average distance, standard deviation indicates that some ROIs are harder to position than others. Regions such as the middle sagittal plane or the ROIs related to the projection pathway (internal capsule, midbrain and medulla) all have an average center of mass distances across raters below 4mm with a small standard deviation (between 2mm and 4mm) (Figure 4). However, regions such as the precentral and postcentral gyri or the temporal stem all have a higher average center of mass distances across raters (around 6mm) and a much larger standard deviation (between 4mm and 8mm). These regions reach translation differences as high as 30mm. Figure 4 shows the distribution of distances for each ROI and how raters disagree on their placement. It is important to mention that in the virtual dissection protocol (Tractostorm 2 - Zenodo), some regions that are not independent (described in relation to others) are therefore subject to similar levels of variability. For example, the centrum semiovale region (central CS) is described as a circular ROI located between the precentral and postcentral gyri.

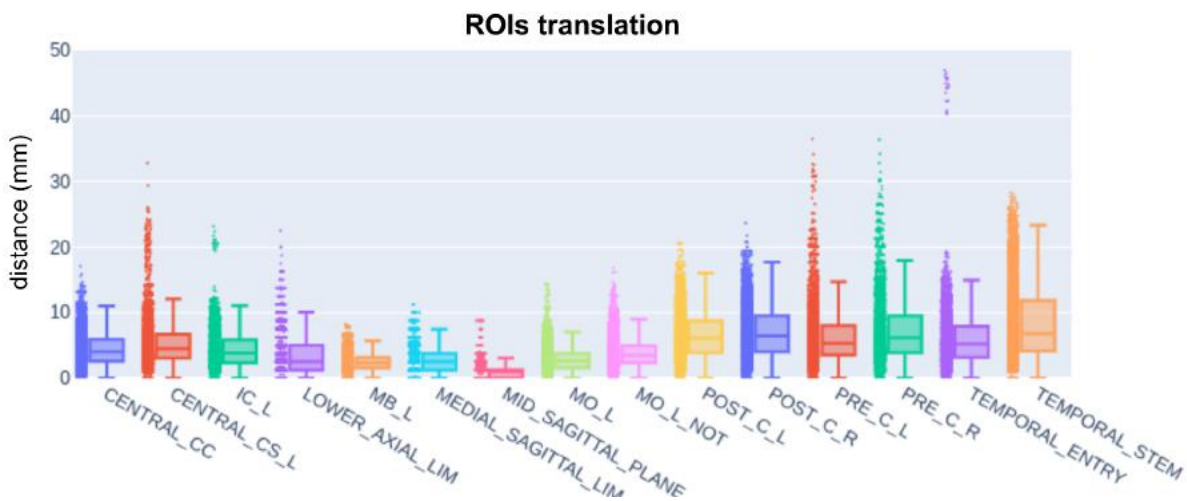

Figure 4: Translation distances (in mm) between raters segmentation for each ROI (center of mass). Low values mean that raters positioned their region closely (e.g. mid-sagittal plane), high values mean that raters did not agree on positioning and placed their ROI far apart (e.g. temporal stem). This can be interpreted as the ease (or difficulty) to locate the slice indicated by the protocol.

Disagreement on ROIs' shape was lower. As a matter of fact, instructions and examples detailing the expected shape constrained the raters to similar delineation. For example, all circular ROIs had a proposed approximate radius in the instructions. This means that the distance shown in Figure 5 for the internal capsule (IC), midbrain (MB), medulla (MO), centrum semiovale (central CS) and temporal stem were simply the distribution of differences in radius between raters. Whole slice ROIs covering the entire plane such as the lower axial limit, the

medial sagittal limit or the middle sagittal plane had little to no shape difference. The outliers seen in Figure 4 for these planar regions were due to raters drawing a very large rectangle rather than filling the whole slice. The ROIs with the highest variability in shape are the large hand-drawn regions, such as the precentral and postcentral gyri (left and right).

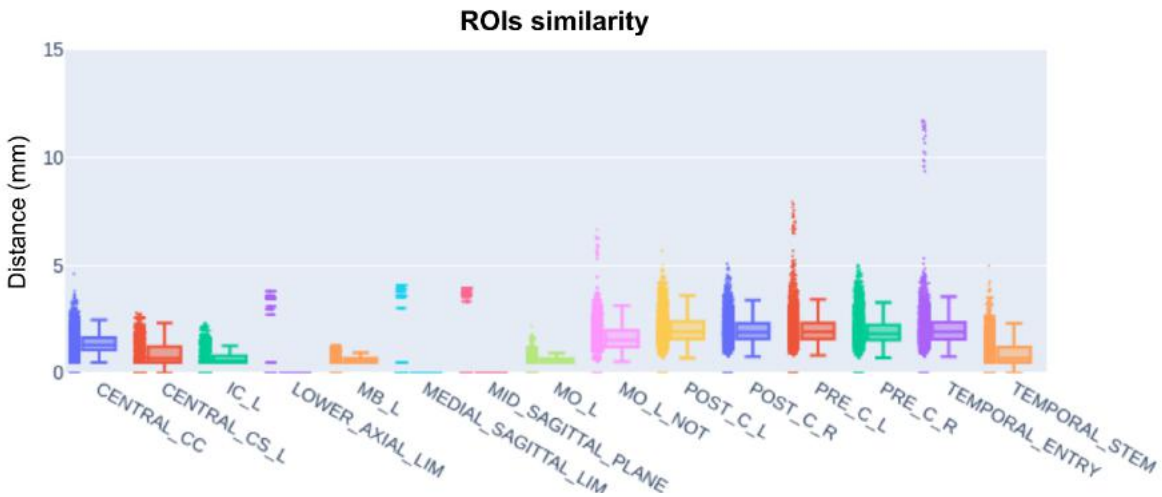

Figure 5: Similarity distances (in mm) between raters segmentation for each ROI (mean Hausdorff distance). Low values mean that raters drew by hand (or circular region) a similar shape (or radius). High values mean that raters disagree on the shape/size of regions. For example, all 3 planar regions are extremely similar in shape because most raters selected the entire slice, thus identical shape. Regions such as the internal capsule or temporal stem are circular and the difference is due to a different choice of radius in the software. This can be interpreted as the ease (or difficulty) to choose the appropriate shape/size of the regions indicated by the protocol.

2.2 Regions of interest impact on pathways

By removing one ROI from the set of rules during a pathway virtual dissection, we can identify if it is useful to the process, however, this is limited to a single removal as the interaction between ROI could have unexpected consequences on the final bundle. In Figure 6, we can see ROIs that are critical to the virtual dissection and others that are not affecting the final bundles. The body of the corpus callosum (central CC) is an example of a crucial ROI, the Dice score between the baseline bundle and the bundle dissected without using the central CC criterion average at 0.25. Without it, the resulting pathway encompasses commissural, association and projection streamlines and lead to a completely different morphology (closer to a whole-brain than a bundle). The lower axial limit, temporal entry and internal capsule (IC) are examples of important ROIs (Dice scores of 0.8, 0.8 and 0.75 respectively). Removing them changes the resulting bundle, but remains somewhat similar to the baseline. The temporal stem, postcentral gyrus (for the arcuate fasciculus), and midbrain (MB) are examples of ROIs with low impact (all Dice scores averaging at 0.99). Removing these ROIs from the set of rules of the virtual dissection protocol does not significantly affect the shape.

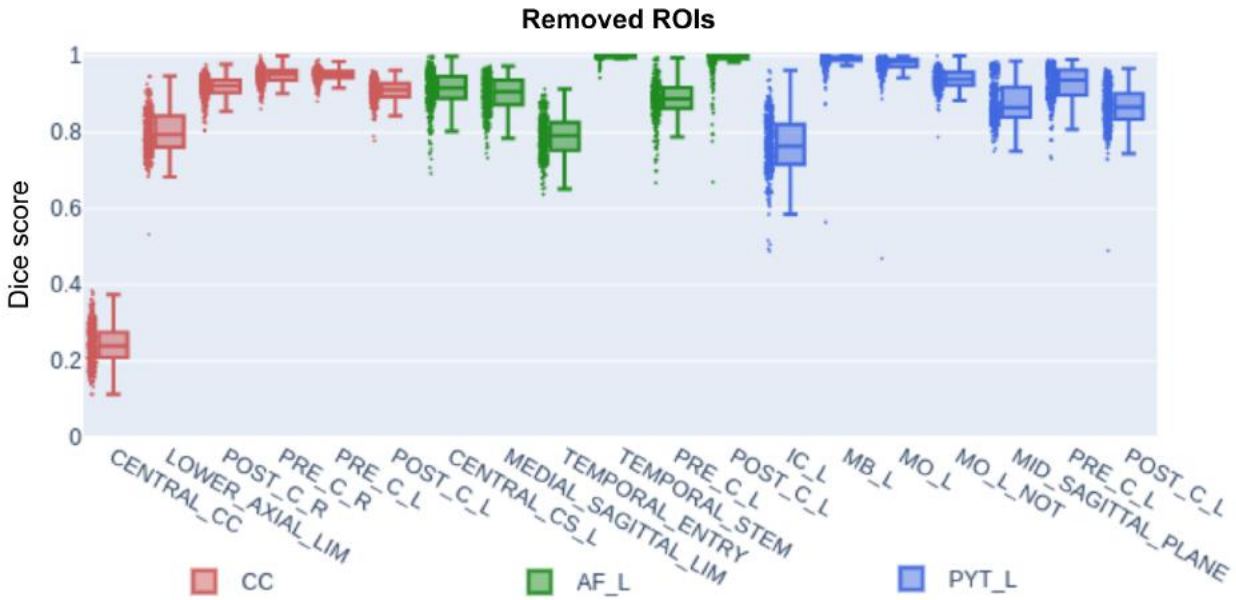

Figure 6: Dice score representing the overlap of bundles (*RGB: CC, AF, PYT*) when 1 ROI is removed from the set of segmentation rules. The ROIs shown on the X-axis are the regions removed from the set of rules. High scores with low variability mean that the ROI impact on the final segmentation is negligible (e.g. temporal stem). Low scores with high variability mean that ROI impact on the final segmentation is substantial (e.g. body of the corpus callosum, CC). This can be interpreted as the impact the ROI/rule has on the virtual dissection (is it needed or not in the protocol)

Identifying the ROIs that do not contribute to the virtual dissection protocol is one way to probe the set of rules, but the variability caused by each ROI is another important insight. The intrinsic variability of ROIs does not directly inform us on the resulting bundle variability. Figure 7 shows the variability caused by each ROI when substituted from all raters. ROIs substitution indicates all ROIs contribute (at least partially) to the overall bundle variability. However, while the average Dice scores are high, some ROIs have a large standard deviation. The central CC, temporal entry, precentral gyrus (of the arcuate fasciculus) and the internal capsule (IC) have standard deviations above 0.1 with Dice scores lower value reaching as low as 0.3.

The variability caused by a ROI present in each set of rules (3 BOIs) is not equal across BOIs, meaning that the interaction with the different rules in the virtual dissection process is likely unique. For example, the precentral gyrus is used as an exclusion criterion in the CC and PYT and both have very low variability and small standard deviations (meaning that raters' variability for that ROI has a low and stable impact on the BOI. However, the same ROI when used as an inclusion criterion for the AF has a much larger standard deviation (meaning that raters' variability is less stable when used with this specific virtual dissection).

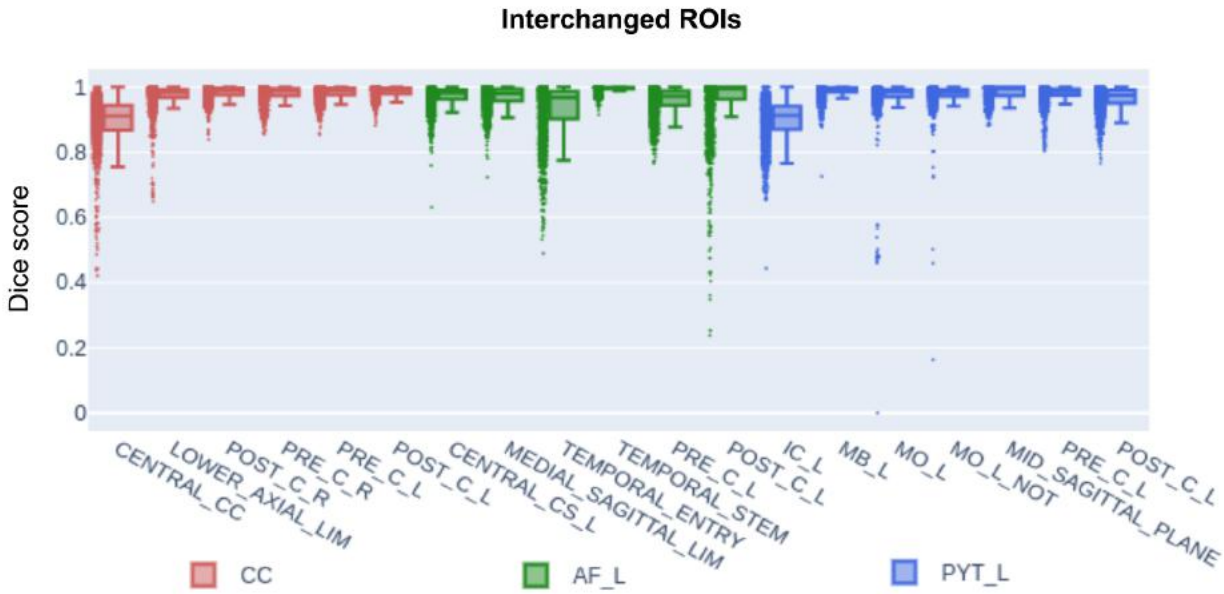

Figure 7: Dice score representing the overlap of bundles (*RGB: CC, AF, PYT*) when 1 ROI from a specific rater is substituted with the same ROI from another rater. High scores with low variability mean that the ROI decisions have a limited impact on the final results (e.g. temporal stem, AF). Low scores with high variability mean that ROI decisions will, sometimes, have a major impact on the resulting bundle (e.g. internal capsule, PYT). This can be interpreted as how much a ROI variability has on the bundle's virtual dissection variability. The red (CC) boxplots are linked to their visual counterpart in Figure 8. Same for the green (AF) with Figure 9 and blue (PYT) with Figure 10.

This ROI-by-ROI effect is explored in more detail in the supplementary materials. This visualization is necessary because it is difficult to choose a quantitative threshold based on the boxplots. A good Dice Score (e.g. >0.9) can be achieved while drastically reducing the anatomical validity of a resulting bundle. For this reason, the qualitative inspection of Figures 8 to 10 is crucial to complement the score of Figure 7. Figure 8 to 10 are previews of the visualization technique used to explore how the ROI variability impacts each bundle in the virtual dissection protocol. The yellow coloring represents the lower value of the reproducibility spectrum. In other words, the more the streamlines' color tends towards yellow, the more they are inconsistently included in the final bundles due to higher variability in the ROI positioning. On the opposite, the more the streamlines' color tends towards dark blue, the more they are included in the bundle. This qualitative visualization provides insight into the spatial impact of each ROI's variability and should be seen as complementary information to Figure 7.

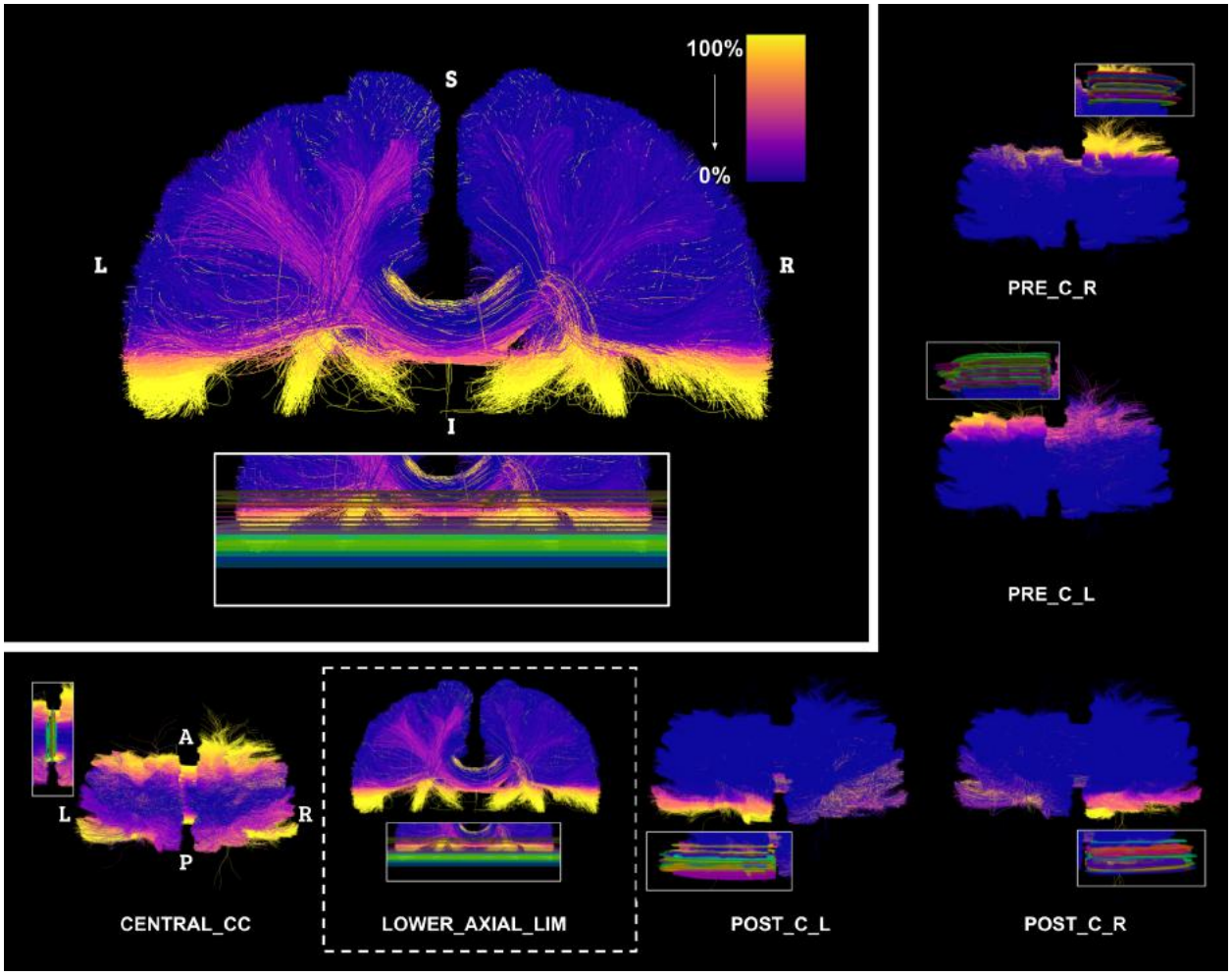

Figure 8: Mosaic of all ROIs involved in the corpus callosum virtual dissection. Streamlines variability, due to ROI placement, is encoded in the coloring. The **lower axial limit** is an interesting example of how differences in positioning can affect the bundle. The goal of this region is to prevent streamlines from entering a projection pathway or reaching into the nuclei (**exclusion** criterion). While the majority of raters selected an axial slice low enough to achieve that goal without impacting the bundle, a few raters selected an axial place that started too high and impacted the lateral fanning in the precentral and postcentral gyri. The instruction could have been made clearer and the ROI modified to simply be a large rectangle that only blocks the most central region of the brain rather than the entire slice. This way, the same goal is achieved without impacting the fanning regions.

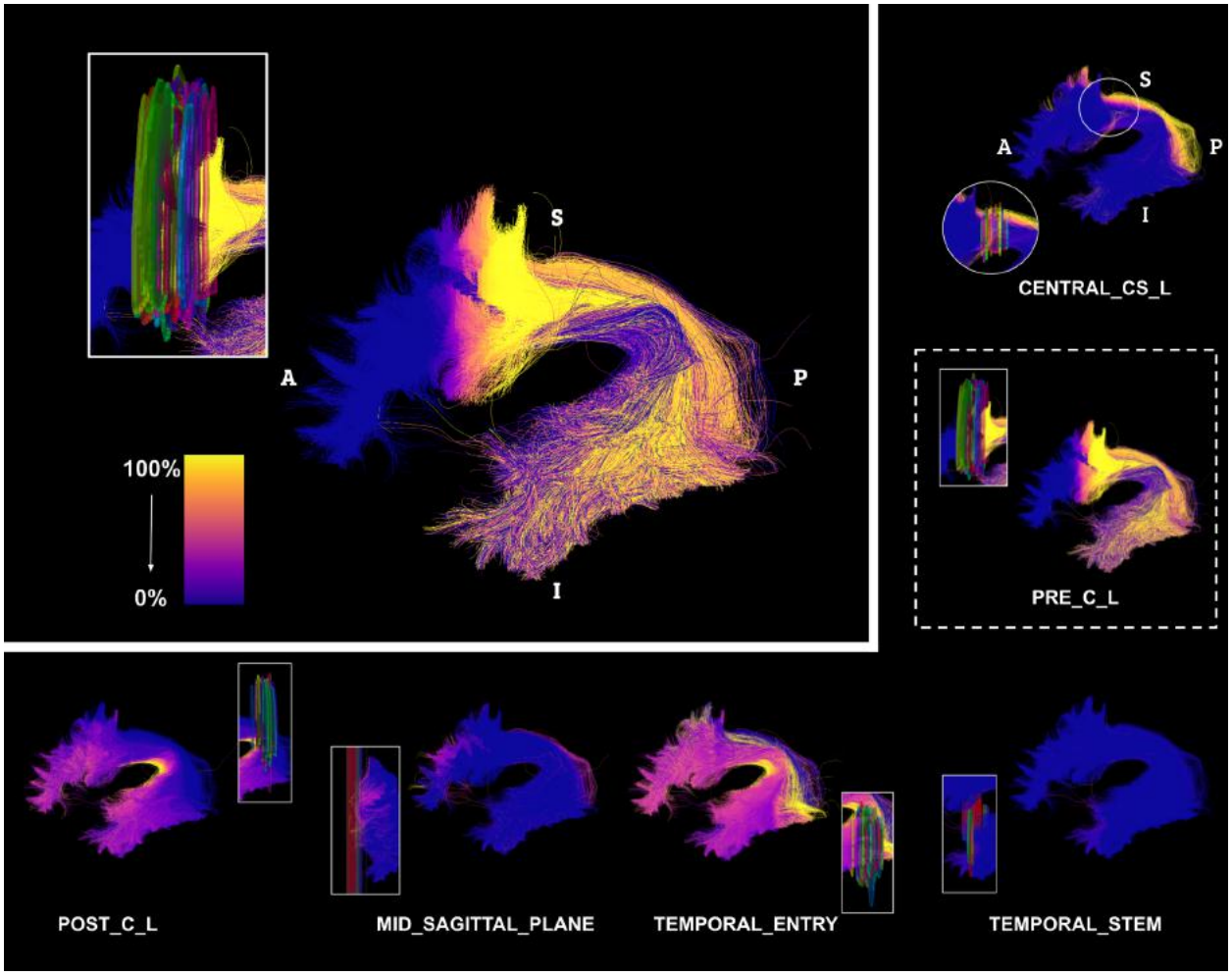

Figure 9: Mosaic of all ROIs involved in the arcuate fasciculus virtual dissection. Streamlines variability, due to ROI placement, is encoded in the coloring. The **left precentral gyrus** is an example of a non-anatomically useful ROI that leads to high variability. As shown in (Vavassori et al., 2021), the AF can reach into the precentral gyrus and this ROI forces the AF to go past the precentral gyrus. The 'rainbow' coloring observable in the figure shows that the difference in positioning was large enough to either include or exclude the precentral gyrus in the bundle. As opposed to Figure 8, this ROI was used as an **inclusion** criterion, meaning that when being positioned too anteriorly the dissection discarded the streamlines not reaching the ROI. By removing the ROI from the set of rules of the virtual dissection of the AF, not only the bundle would be more anatomically accurate, but a large source of variability would be removed.

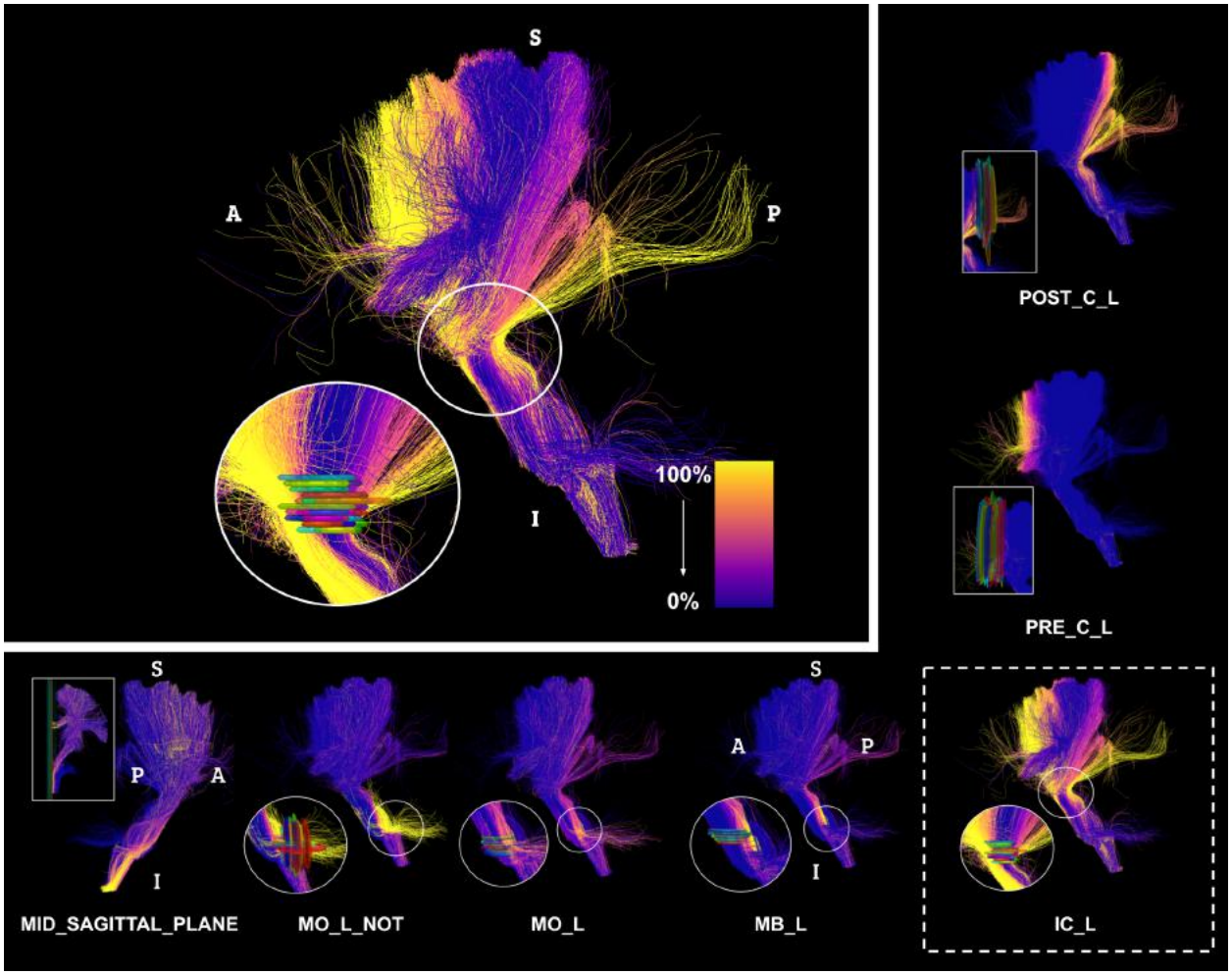

Figure 10: Mosaic of all ROIs involved in the pyramidal tract virtual dissection. Streamlines variability, due to ROI placement, is encoded in the coloring. The **internal capsule** is crucial to isolate the pyramidal tract as it is a major bottleneck. However, the difference in positioning and the ROIs' radii leads to large variation (streamlines reaching too anteriorly and/or posteriorly). Positioning disagreement in the inferior/superior axis does not have a major impact, this axis is parallel to the global orientation of the pathway close to the internal capsule. Disagreement in the positioning only has a major impact on variability if that disagreement is perpendicular to the global orientation of the pathways (anterior/posterior or left/right). While not exactly anatomically justified, enlarging the ROI to guarantee that it will most likely encompass the whole internal capsule would decrease the risk of variability.

3. Discussion

From previous projects we learned that a protocol can be improved with an online course (Rheault et al., 2022b), we learned that raters are a bigger source of variability than acquisition schemes or data quality (Rheault et al., 2022a), we learned that with appropriate protocol experts and nonexperts have similar levels of reproducibility (Rheault et al., 2020a), we learned that there are ways to quantify the source of variability and (iteratively) improve a virtual dissection protocol and that ROIs must be individually evaluated in a bundle-wise manner to truly capture their impact and contribution (this work). These projects demonstrated that virtual dissection protocol cannot be generalized when changing or introducing new variables. The differences between the protocols used for the previously mentioned projects (software, tractography algorithms, teaching methods, etc.) led to heterogeneous results for anatomically similar bundles. These efforts put toward the investigation of WM virtual dissection expose the need for better protocols and anatomical definitions with more rigorous instructions. The following section explores ways the protocol could be optimized to reduce variability or decrease the amount of work from raters. This shows the importance of having an iterative process where decisions are not only justified but tested in "*the field*" to adjust a protocol to reach the desired anatomical accuracy, reproducibility and workload.

3.1 Regions of interest contributions to the virtual dissection protocol

This analysis, which is focused on ROIs, allowed us to determine each ROI intrinsic variability and the impact it has when used in the virtual dissection protocol. The ROIs with the highest positional variability are the **temporal stem region**. Multiple factors explain why this ROI has such a high variability. First, the starting approximate position for the ROI only refers to the Sylvian fissure anterior to the precentral gyrus ROI (another region with high variability), this leaves the raters to pick a sagittal slice to identify the Sylvian fissure, then since it is an elongated landmark more than one position is appropriate. The rest of the instruction involves a green hypersignal in the temporal lobe (stem of the arcuate fasciculus), this landmark can be observed in multiple slices.
Providing more details could reduce variability, but this leads to the next important point: This variability has little to no effect on the virtual dissection of the arcuate fasciculus. Not only is this ROI difficult to position and has high variability, but it has no contribution to the dissection process. This ROI can be removed from the AF set of rules.

The precentral and postcentral gyri (and the central centrum semiovale, defined using the precentral and postcentral gyri) show high variability too. The precentral gyri have a higher variability for similar reasons to the temporal stem as the initial approximate position involves choosing a sagittal slice to identify the Sylvian fissure and then identify the precentral gyrus (generally along the inferior/superior axis). Despite its high variability, this ROI is useful to the pyramidal tract and corpus callosum virtual dissection process (as an exclusion rule) and high intrinsic variability only has a small impact on the bundles. However, the same ROI is used (as an inclusion) for the arcuate fasciculus and in this case, the high intrinsic variability does translate to high virtual dissection variability.

In this case, the ROI is anatomically unnecessary (The ROI being anterior to the precentral when in fact the AF should fan into the precentral) and generate high variability. This ROI should be removed from the AF set of rules, but kept for both the CC and PYT.

Since there are two recommendations for the AF set of rules (removing the temporal stem and the precentral gyrus ROIs), we verified that applying the new protocol for the AF did not lead to unforeseen results. This could be due to streamlines (that needed both conditions) not being discarded and decreasing reproducibility or anatomical validity. The results were as expected, the removal of the precentral gyrus condition decreased variability and led to an anatomically plausible bundle and the removal of the temporal stem condition did not lead to a significant increase of spurious streamlines. Therefore, both conditions can safely be removed from the protocol.

More detailed analyses and recommendations for each ROI are available in the supplementary materials sections. To summarize, there is a balance between 3 spectrums: Anatomical meaning, effect on the virtual dissection and difficulty to delineate. There is no perfect way to quantify these criteria, anatomical justification must emerge from prior knowledge in neuroanatomy and agreement between experts to justify or not the use of a specific region of interest. However, we believe that the approach used in Figures 4 and 5 helped us inform the “difficulty to delineate” spectrum and Figures 6 and 7 helped us inform the “Effect on dissection” spectrum. The difficulty to delineate applies to each ROI individually, but the anatomical justification and the effect on dissection will vary from bundle to bundle (in which the ROI is used in the set of criteria). Figure 11 summarizes how we evaluated each ROI in the supplementary material to reach useful recommendations to adapt the protocol.

Sadly, this balance cannot be strict and fixed in time. A virtual dissection protocol targeting neurosurgery applications will not have the same requirements as a protocol design for thousands of healthy subjects, for example. The difficulty to delineate can be different if the raters are experts in neuroanatomy versus nonexperts. However, we believe that the framework presented in this study is required to be able to rank (even if only qualitatively) each ROI on these spectrums and acknowledge the balance to justify their advantage and inconvenience to the protocol.

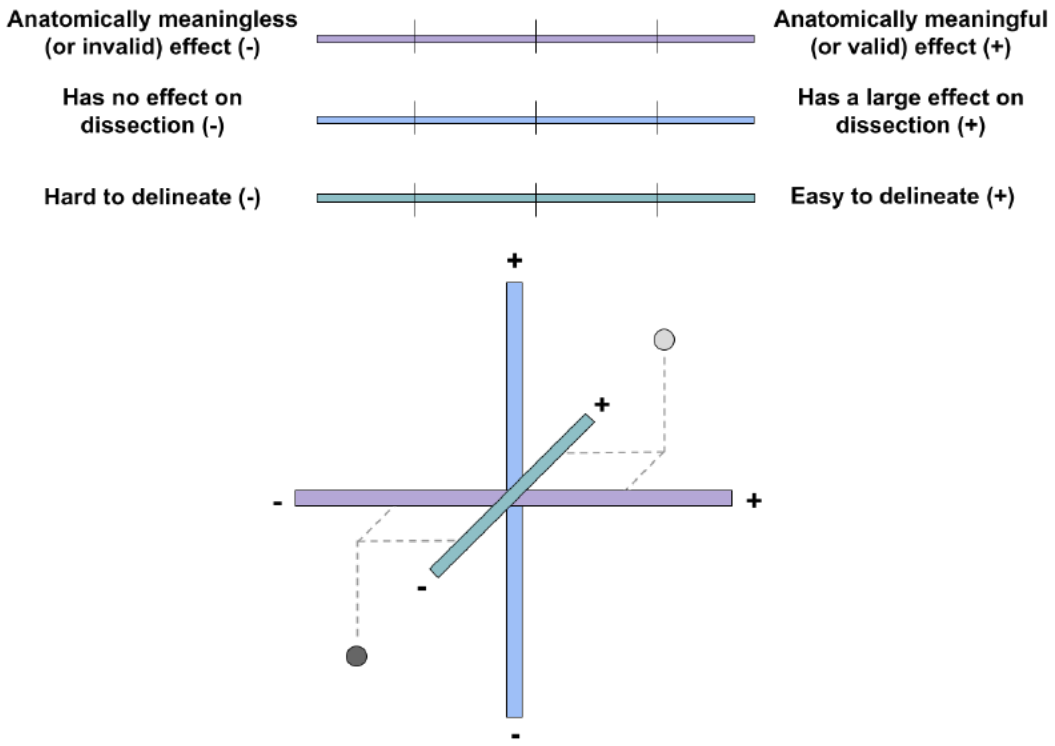

Figure 11: In the current study 3 main spectrums describe each ROI in the virtual dissection protocol. While mostly a qualitative ranking, it is important to consider all 3 when providing recommendations. The extreme cases are easy, such as a ROI easy to delineate, with large and meaningful/valid effect on the dissection (and the opposite). ROI with no effect or invalid effect should be removed, ROIs that are hard to delineate with large effect should have their instructions reviewed to reduce variability.

3.2 General recommendations

From the multiple projects (including this one) related to practical definitions and virtual dissection reproducibility involving the authors of this project (Maier-Hein et al., 2017; Chenot et al., 2019; Bayrak et al., 2019; Rheault et al., 2020a; Nath et al., 2020; Schilling et al., 2021a, 2021b; Rheault et al., 2022a, 2022b) or from other groups (Wakana et al., 2007; Catani and De Schotten, 2008; Wassermann et al., 2016; Van Hecke et al., 2016; Warrington et al., 2020; Maffei et al., 2021) led us to propose some recommendations for the definition of WM pathways' practical extraction. These recommendations are also inspired by guidelines of the EADC-ADNI Harmonized Protocol for manual hippocampal segmentation (Frisoni et al., 2015). These general guidelines are provided to help future protocols design, some adaptations could be required for young children, older adults, lesioned brains. However, it is likely that virtual dissection of such datasets would need a targeted protocol with instructions tailored to the population and that these general recommendations would still be relevant for its design. More details for each recommendation are available in the supplementary materials.

Dissection in standardized space. Virtual dissection protocol should always be performed in a standardized space (e.g. MNI or AC-PC aligned). The reason for this recommendation is simple: Landmarks do not look the same depending on the brain orientation due to slicing. Relative spatial location is easier to interpret (by human or by computer) in MNI/AC-PC align space (medial to, lateral to).

Clear instruction on the software does help raters. While the description of regions of interest is crucial, a detailed presentation of the software is likely as important. Presenting functionalities, tips to prevent major pitfalls, avoiding crashes and saving time all are important to reduce the mental workload of a rater. Written text is a good medium for a reference guide, but step-by-step instruction (or better, a video) is a good way to increase reproducibility.

Virtual dissection workflow. Another way to reduce variability is to establish a workflow for the tasks. Organizing the input, the order in which the data should be loaded in the software, and requiring a naming convention for all ROIs and BOIs are all solutions to reduce the workload of raters.

Getting in the zone. While the accuracy of a virtual dissection does not significantly change over time, each rater workflow (software manipulation, files structure, details about the protocol, filename convention) needs some time to get up to speed. It is good practice to set time to focus and perform a few virtual dissections at the time and to always complete a dataset before stopping.

Being too strict is worse than being permissive. Using very small ROIs or using too many exclusion regions is not only harder for raters, but it leads to poor results. Using permissive and easier to locate ROIs leads to good results. Combining multiple ROIs will facilitate the process of virtual dissection. A good balance of inclusion and exclusion is a better approach than relying mostly on one or the other, but using too many ROIs can increase the workload of raters.

Using other ROIs as landmarks. This is generally a good practice to facilitate and accelerate the virtual dissection process. By providing a lot of details on specific ROIs, the subsequent regions should become easier to locate as they reference previously drawn ones. However, this is true only if used for ROIs that are highly consistent. A region with high variability will ‘transfer’ its variability to any other region referencing it.

Generating more than one ROI from the same landmarks. Only re-use ROIs if they have the same impact, if not a new ROI may be needed. For example, if a ROI is used twice, once as inclusion and once as exclusion, and they lead to different effects on their respective dissection, it could be more reproducible and easier for raters to create two separate ROIs from the same set of instructions.

Providing ranges. Each ROIs should have an approximate initial location and an approximate size to accommodate the anatomical variability of subjects. When in a standardized space, a range of slices can be provided to facilitate location. The exact value should be avoided, raters should be instructed to start from an approximate location and scroll until something specific is located.

Locating landmarks. It appears that instructions using the RGB map are easier to follow than the ones using only T1. This is likely because RGB contains hypersignal of 3 main colors, much easier to describe and locate than a general area of 'gray' on the T1. It also appears that descriptions that involve multiple orientations at once are difficult (axial, coronal sagittal). Landmarks that are perpendicular to the slicing orientation should be prioritized to limit switching from one plane to the other (in the software).

Optimal ROIs design. When using the **inclusion rule** for a stem/core, the positioning along the pathway is less important, the focus should be targeted at the perpendicular positioning (perpendicular to a strongly oriented core). A larger diameter is likely better than a smaller one. When using the **exclusion rule**, bigger is often better as long as you do not accidentally cut in half the pathway somewhere else (e.g. middle sagittal plane to prevent commissural streamlines accidentally blocking the brainstem). For exclusion rules, adding a few millimeters in the opposite direction of what is supposed to be stopped provides a good margin of error to reduce variability.

Increasing complexity. The virtual dissection protocol should start with basic descriptions of the bundles of interests, a summary of the steps that will follow and a general overview of all regions that will be drawn. ROIs' instruction should start with a fast (and less accurate) set of instructions to delineate the region followed by a slower (more accurate) detailed instruction.

Cleaning ROIs. If cleaning ROIs are needed (not anatomically meaningful, but simply to clean outliers from tractography limitations), they should be described and justified, follow the previously listed recommendations and all cleaning ROIs should be accounted for. It should not be left to the raters to decide how big, how many, where cleaning ROIs are needed.

3.3 Transfer of knowledge: Why are these projects and recommendations useful? (Opinions from co-authors)

For a neuroanatomist?

More than any other protagonist of the current study, the neuroanatomist faces the crucial question of the definition and denomination of a given white matter pathway, especially when confronted with the question of its virtual dissection in tractography (Mandonnet et al., 2018; Panesar and Fernandez-Miranda, 2019; Porto de Oliveira et al., 2021; Vavassori et al., 2021). There is a serious lack of consensus on how to describe the fascicular organization of the white matter of the human brain, mainly because we do not yet understand it well enough. A commonly accepted definition of a given fascicle will only be possible by multiplying the definition criteria. The Tractostorm projects and the resulting recommendations show, both

qualitatively and quantitatively, how the number and position of ROIs that will characterize part or all of a bundle's path through white matter must be part of this consensus definition in the future. The process needed to reach such a consensus, likely through iteration, will likely need tractography and virtual dissection.

For a neurosurgeon?

The results of the current work highlight the difficulty in achieving optimal intra-rater and inter-rater reproducibility, which may in part contribute to the discordant findings in neurosurgical studies that report on the utility of tractography in the perioperative management of brain pathologies. Understanding these drawbacks is primordial to correctly interpret the results of studies involving bundle dissection and characterization. Nevertheless, the effort made by these projects in establishing reproducible dissections opens the door to a framework that has the potential to shift the currently imperfect paradigm towards a unifying methodical approach and prompt identification of WM bundles.

Despite the studied protocols targeting healthy subjects, the current series of studies have direct implications in the selection, presurgical evaluation, operative and postoperative management of patients affected by various neurological diseases. Being able to accurately characterize WM tracts has the potential to directly influence patient care. A better understanding of WM pathways segmentation (even in healthy subjects) is a step forward. Simple, accurate and reproducible definitions of ROIs and concomitant bundles are essential to 1) establish correlations between virtually dissected WM tracts and clinical symptoms, 2) perform longitudinal studies of brain plasticity and 3) compare and aggregate results of different clinical trials. Ultimately, this work led to a better understanding of how tractography can be interpreted (with all of its challenges and biases in virtual dissection).

For a data scientist?

The results of this work are immediately useful for data scientists. A data scientist may be one who ‘creates’ the tools - i.e., the software, algorithms, and dissection protocols - or one who ‘utilizes’ these tools and processes. For those who create the tools, the process of targeting a pathway must be supplemented with clear instructions, with emphasis on the steps that are more critical for reduced anatomical variability. It is necessary to understand sources of variability to reduce it, which will ultimately lead to the identification of smaller effect sizes when studying changes across pathways, subjects, or time. Not only can these results be used to continually improve manual protocols, but also used for creating automated or semi-automatic segmentation protocols. Methods based on regions in a standard space, or those developed to learn streamline or region placement, must take into account the lessons learned here to enable robust and reproducible segmentation and subsequent analysis. Training and validation of semi- or fully-automated algorithms must consider the variability inherent in generating the ‘ground truth’ bundles upon which it is trained and compared against. Finally, those that utilize the tools must be aware of variance due to intra/inter-sites, intra/inter-raters, and variability inherent within the methodology to draw informed conclusions about the studied pathways.

4. Conclusion

This work shows what is driving differences between virtual dissection. Our analysis shows that despite high variability, manual delineation can have, at the same time, low impact on some bundles and high impact on others. These results demonstrate the importance of not only quantifying variability but that it must be done for every combination of regions of interest and pathways.

In this study, we propose a framework to evaluate the individual contribution of regions of interest in a virtual dissection protocol. This information is complementary to prior work from the same group of authors and contributes to better design anatomical definitions (practical or theoretical) of white matter pathways. Our in-depth investigation of the source of variability provides insight into more efficient ways to transmit knowledge related to structural neuroanatomy and tractography.

Our current and prior contributions were incremental and tedious and there are many technical challenges still to be resolved, but we paved the way forward to a more coherent approach to convey knowledge in neuroanatomy and tractography. Having more efficient communication between groups with different expertise is crucial in a multidisciplinary field and we hope that our work brings us closer to a consensus on how to define major white matter pathways and reproducible/replicable science.

Funding

This work was supported by the National Institutes of Health (NIH) under award numbers R01EB017230, P50HD103537, T32EB001628, and in part by ViSE/VICTR VR3029 and the National Center for Research Resources, Grant UL1 RR024975-01. We thank the Universite de Sherbrooke institutional research chair in Neuroinformatics that supports Maxime Descoteaux and his team. Sami Obaid was supported by the Savoy Foundation studentship and from scholarships from the Fonds de Recherche du Québec - Santé (277581)

Data Availability

This study has been reviewed by the Internal Review Board (IRB) of Vanderbilt University (#211156). All raters agreed to collaborates for the data gathering phase. The datasets generated before the study and the obtained files from the raters (analyzed during the current study) are available on Zenodo: https://zenodo.org/record/5190145#.YhZornXMKiM

Bibliography

Agrawal, A., Kapfhammer, J.P., Kress, A., Wichers, H., Deep, A., Feindel, W., Sonntag, V.K., Spetzler, R.F., Preul, M.C., 2011. Josef Klingler's models of white matter tracts: influences on neuroanatomy, neurosurgery, and neuroimaging. Neurosurgery 69, 238–254.

Alturkistani, H.A., Tashkandi, F.M., Mohammedsaleh, Z.M., 2016. Histological stains: a literature review and case study. Glob. J. Health Sci. 8, 72.

Arantes, M., Arantes, J., Ferreira, M., 2018. Tools and resources for neuroanatomy education: a systematic review. BMC Med. Educ. 18, 1–15.

Axer, M., Strohmer, S., Gräßel, D., Bücker, O., Dohmen, M., Reckfort, J., Zilles, K., Amunts, K., 2016. Estimating fiber orientation distribution functions in 3D-polarized light imaging. Front. Neuroanat. 10, 40.

Bayrak, R.G., Schilling, K.G., Greer, J.M., Hansen, C.B., Greer, C.M., Blaber, J.A., Williams, O., Beason-Held, L.L., Resnick, S.M., Rogers, B.P., others, 2019. TractEM: Fast Protocols for Whole Brain Deterministic Tractography-Based White Matter Atlas. bioRxiv 651935.

Benedictis, A., Petit, L., Descoteaux, M., Marras, C.E., Barbareschi, M., Corsini, F., Dallabona, M., Chioffi, F., Sarubbo, S., 2016. New insights in the homotopic and heterotopic connectivity of the frontal portion of the human corpus callosum revealed by microdissection and diffusion tractography. Hum. Brain Mapp. 37, 4718–4735.

Catani, M., De Schotten, M.T., 2008. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. cortex 44, 1105–1132.

Chang, B.S., Molnár, Z., 2015. Practical neuroanatomy teaching in the 21st century. Ann. Neurol. 77, 911–916.

Chenot, Q., Tzourio-Mazoyer, N., Rheault, F., Descoteaux, M., Crivello, F., Zago, L., Mellet, E., Jobard, G., Joliot, M., Mazoyer, B., others, 2019. A population-based atlas of the human pyramidal tract in 410 healthy participants. Brain Struct. Funct. 224, 599–612.

Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis: I. Segmentation and surface reconstruction. Neuroimage 9, 179–194.

David, S., Heemskerk, A.M., Corrivetti, F., Thiebaut De Schotten, M., Sarubbo, S., Corsini, F., De Benedictis, A., Petit, L., Viergever, M.A., Jones, D.K., others, 2019. The superoanterior fasciculus (SAF): A novel white matter pathway in the human brain? Front. Neuroanat. 13, 24.

Dick, A.S., Tremblay, P., 2012. Beyond the arcuate fasciculus: consensus and controversy in the connectional anatomy of language. Brain 135, 3529–3550.

Fernández-Miranda, J.C., Wang, Y., Pathak, S., Stefaneau, L., Verstynen, T., Yeh, F.-C., 2015. Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. Brain Struct. Funct. 220, 1665–1680.

Frisoni, G.B., Jack Jr, C.R., Bocchetta, M., Bauer, C., Frederiksen, K.S., Liu, Y., Preboske, G., Swihart, T., Blair, M., Cavedo, E., others, 2015. The EADC-ADNI Harmonized Protocol for manual hippocampal segmentation on magnetic resonance: evidence of validity. Alzheimers Dement. 11, 111–125.

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., 2014. Dipy, a library for the analysis of diffusion MRI data. Front. Neuroinformatics 8, 8.

Garyfallidis, E., Côté, M.-A., Rheault, F., Sidhu, J., Hau, J., Petit, L., Fortin, D., Cunanne, S., Descoteaux, M., 2018. Recognition of white matter bundles using local and global streamline-based registration and clustering. NeuroImage 170, 283–295.

Hau, J., Sarubbo, S., Houde, J.C., Corsini, F., Girard, G., Deledalle, C., Crivello, F., Zago, L., Mellet, E., Jobard, G., others, 2017. Revisiting the human uncinate fasciculus, its subcomponents and asymmetries with stem-based tractography and microdissection

validation. Brain Struct. Funct. 222, 1645–1662.
Huttenlocher, D.P., Klanderman, G.A., Rucklidge, W.J., 1993. Comparing images using the Hausdorff distance. IEEE Trans. Pattern Anal. Mach. Intell. 15, 850–863.
Jbabdi, S., Johansen-Berg, H., 2011. Tractography: where do we go from here? Brain Connect. 1, 169–183.
Maffei, C., Jovicich, J., De Benedictis, A., Corsini, F., Barbareschi, M., Chioffi, F., Sarubbo, S., 2018. Topography of the human acoustic radiation as revealed by ex vivo fibers micro-dissection and in vivo diffusion-based tractography. Brain Struct. Funct. 223, 449–459.
Maffei, C., Lee, C., Planich, M., Ramprasad, M., Ravi, N., Trainor, D., Urban, Z., Kim, M., Jones, R., Henin, A., others, 2021. Using diffusion MRI data acquired with ultra-high gradients to improve tractography in routine-quality data. bioRxiv.
Maier-Hein, K.H., Neher, P.F., Houde, J.-C., Côté, M.-A., Garyfallidis, E., Zhong, J., Chamberland, M., Yeh, F.-C., Lin, Y.-C., Ji, Q., others, 2017. The challenge of mapping the human connectome based on diffusion tractography. Nat. Commun. 8, 1349.
Mandonnet, E., Sarubbo, S., Petit, L., 2018. The nomenclature of human white matter association pathways: Proposal for a systematic taxonomic anatomical classification. Front. Neuroanat. 12, 94.
Nath, V., Schilling, K.G., Parvathaneni, P., Huo, Y., Blaber, J.A., Hainline, A.E., Barakovic, M., Romascano, D., Rafael-Patino, J., Frigo, M., others, 2020. Tractography reproducibility challenge with empirical data (TraCED): the 2017 ISMRM diffusion study group challenge. J. Magn. Reson. Imaging 51, 234–249.
O'Donnell, L.J., Pasternak, O., 2015. Does diffusion MRI tell us anything about the white matter? An overview of methods and pitfalls. Schizophr. Res. 161, 133–141.
Panesar, S., Fernandez-Miranda, J., 2019. Commentary: The Nomenclature of Human White Matter Association Pathways: Proposal for a Systematic Taxonomic Anatomical Classification. Front. Neuroanat. 13, 61.
Porto de Oliveira, J.V.M., Raquelo-Menegassio, A.F., Maldonado, I.L., 2021. What's your name again? A review of the superior longitudinal and arcuate fasciculus evolving nomenclature. Clin. Anat. 34, 1101–1110.
Poupon, C., Rieul, B., Kezele, I., Perrin, M., Poupon, F., Mangin, J.-F., 2008. New diffusion phantoms dedicated to the study and validation of high-angular-resolution diffusion imaging (HARDI) models. Magn. Reson. Med. Off. J. Int. Soc. Magn. Reson. Med. 60, 1276–1283.
Rheault, F., Bayrak, R.G., Wang, X., Schilling, K.G., Greer, J.M., Hansen, C.B., Kerley, C., Ramadass, K., Remedios, L.W., Blaber, J.A., others, 2022a. TractEM: Evaluation of protocols for deterministic tractography white matter atlas. Magn. Reson. Imaging 85, 44–56.
Rheault, F., De Benedictis, A., Daducci, A., Maffei, C., Tax, C.M., Romascano, D., Caverzasi, E., Morency, F.C., Corrivetti, F., Pestilli, F., others, 2020a. Tractostorm: The what, why, and how of tractography dissection reproducibility. Hum. Brain Mapp.
Rheault, F., Houde, J.-C., Goyette, N., Morency, F., Descoteaux, M., 2016. MI-Brain, a software to handle tractograms and perform interactive virtual dissection, in: ISMRM Diffusion Study Group Workshop, Lisbon.
Rheault, F., Poulin, P., Caron, A.V., St-Onge, E., Descoteaux, M., 2020b. Common misconceptions, hidden biases and modern challenges of dMRI tractography. J. Neural Eng. 17, 011001.
Rheault, F., Schilling, K.G., Valcourt-Caron, A., Théberge, A., Poirier, C., Grenier, G., Guberman, G.I., Begnoche, J., Legarreta, J.H., Y Cai, L., others, 2022b. Tractostorm 2: Optimizing tractography dissection reproducibility with segmentation protocol dissemination. Hum. Brain Mapp.

Saleeba, C., Dempsey, B.R., Le, S., Goodchild, A.K., McMullan, S., 2019. A student's guide to neural circuit tracing. Front. Neurosci. 13, 897.

Schilling, K.G., Petit, L., Rheault, F., Remedios, S., Pierpaoli, C., Anderson, A.W., Landman, B.A., Descoteaux, M., 2020. Brain connections derived from diffusion MRI tractography can be highly anatomically accurate—if we know where white matter pathways start, where they end, and where they do not go. Brain Struct. Funct. 225, 2387–2402.

Schilling, K.G., Rheault, F., Petit, L., Hansen, C.B., Nath, V., Yeh, F.-C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., others, 2021a. Tractography dissection variability: what happens when 42 groups dissect 14 white matter bundles on the same dataset? NeuroImage 118502.

Schilling, K.G., Tax, C.M., Rheault, F., Hansen, C., Yang, Q., Yeh, F.-C., Cai, L., Anderson, A.W., Landman, B.A., 2021b. Fiber tractography bundle segmentation depends on scanner effects, vendor effects, acquisition resolution, diffusion sampling scheme, diffusion sensitization, and bundle segmentation workflow. NeuroImage 118451.

Thomas, C., Frank, Q.Y., Irfanoglu, M.O., Modi, P., Saleem, K.S., Leopold, D.A., Pierpaoli, C., 2014. Anatomical accuracy of brain connections derived from diffusion MRI tractography is inherently limited. Proc. Natl. Acad. Sci. 111, 16574–16579.

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., Connelly, A., 2019. MRtrix3: A fast, flexible and open software framework for medical image processing and visualization. Neuroimage 202, 116137.

Van Hecke, W., Emsell, L., Sunaert, S., 2016. Diffusion tensor imaging: a practical handbook. Springer.

Vavassori, L., Sarubbo, S., Petit, L., 2021. Hodology of the superior longitudinal system of the human brain: a historical perspective, the current controversies, and a proposal. Brain Struct. Funct. 1–22.

Wakana, S., Caprihan, A., Panzenboeck, M.M., Fallon, J.H., Perry, M., Gollub, R.L., Hua, K., Zhang, J., Jiang, H., Dubey, P., others, 2007. Reproducibility of quantitative tractography methods applied to cerebral white matter. Neuroimage 36, 630–644.

Wang, H., Lenglet, C., Akkin, T., 2015. Structure tensor analysis of serial optical coherence scanner images for mapping fiber orientations and tractography in the brain. J. Biomed. Opt. 20, 036003.

Wang, R., Benner, T., Sorensen, A.G., Wedeen, V.J., 2007. Diffusion toolkit: a software package for diffusion imaging data processing and tractography, in: Proc Intl Soc Mag Reson Med. Berlin.

Warrington, S., Bryant, K.L., Khrapitchev, A.A., Sallet, J., Charquero-Ballester, M., Douaud, G., Jbabdi, S., Mars, R.B., Sotiropoulos, S.N., 2020. XTRACT-Standardized protocols for automated tractography in the human and macaque brain. NeuroImage 116923.

Wassermann, D., Makris, N., Rathi, Y., Shenton, M., Kikinis, R., Kubicki, M., Westin, C.-F., 2016. The white matter query language: a novel approach for describing human white matter anatomy. Brain Struct. Funct. 221, 4705–4721.

Yeh, F., 2021. DSI Studio. Zenodo. https://doi.org/10.5281/zenodo.4978980

Click here to access/download
Electronic Supplementary Material
tractostorm_v2_ROIs_supplementary_materials_BSAF.pdf