

A comparison of diffusion MRI presurgical tractography techniques with intraoperative mapping-based validation

Radwan, A.M.^{1,2}, Emsell, L.^{1,2,3,4}, Vansteelandt, K.^{2,3,4}, Cleeren, E.⁵, Peeters, R.⁶, De Vleeschouwer, S.^{2,7,8}, Theys, T.^{2,7,8}, Dupont P.^{2,9}, Sunaert S.^{1,2,6}

1. KU Leuven, Department of Imaging and pathology, Translational MRI, Leuven, Belgium
2. KU Leuven, Leuven Brain Institute (LBI), Department of Neurosciences, Leuven, Belgium
3. KU Leuven, Department of Neurosciences, Neuropsychiatry, Leuven, Belgium
4. KU Leuven, Department of Geriatric Psychiatry, University Psychiatric Center (UPC), Leuven, Belgium
5. UZ Leuven, Department of Neurology, Leuven, Belgium
6. UZ Leuven, Department of Radiology, Leuven, Belgium
7. KU Leuven, Department of Neurosciences, Research Group Experimental Neurosurgery and Neuroanatomy, Leuven, Belgium
8. UZ Leuven, Department of Neurosurgery, Leuven, Belgium
9. KU Leuven, Laboratory for Cognitive Neurology, Department of Neurosciences, Leuven, Belgium

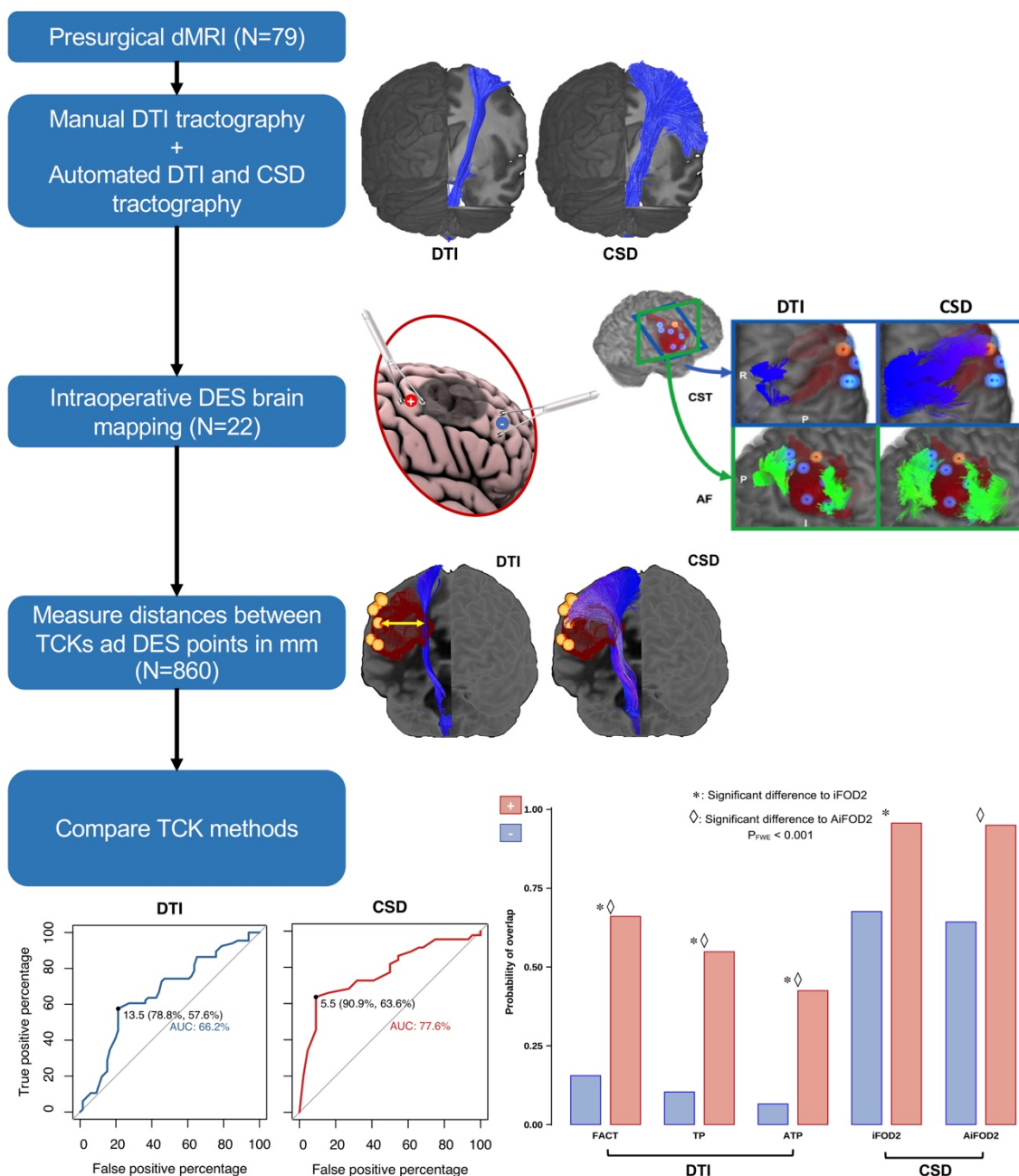
Corresponding author:

Ahmed Radwan, MD, PhD

Email: ahmed.radwan@kuleuven.be

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

A comparison of dMRI pre-surgical tractography techniques with intraoperative mapping-based validation



Graphical abstract: Presurgical white matter mapping using probabilistic CSD tractography is more accurate and sensitive than manual DTI FACT or automated probabilistic DTI tractography. This study included 22 patients with DES data, which was used as the ground truth. Distance in mm between tractograms and DES data resulted in 860 datapoints, 685 of which belonged to the CST and were used for linear modeling, DTI = diffusion tensor imaging, CSD = constrained spherical deconvolution, TCK = tractogram/tractography, FWE = family-wise error rate, AUC = area under the curve

Abstract

Objectives

Accurate presurgical brain mapping enables preoperative risk assessment and intraoperative guidance. This work investigated whether constrained spherical deconvolution (CSD) methods were more accurate than diffusion tensor imaging (DTI)-based methods for presurgical white matter mapping using intraoperative direct electrical stimulation (DES) as the ground truth.

Material and methods

Five different tractography methods were compared (3 DTI-based and 2 CSD-based) in 22 preoperative neurosurgical patients. The corticospinal tract (CST, N=20) and arcuate fasciculus (AF, N=7) bundles were reconstructed, then minimum distances between tractograms and DES coordinates were compared between tractography methods. Receiver-operating characteristic (ROC) curves were used for both bundles. For the CST, binary agreement, linear modeling, and posthoc testing were used to compare tractography methods while correcting for relative lesion and bundle volumes.

Results

Distance measures between 154 positive (functional response, pDES) and negative (no response, nDES) coordinates, and 134 tractograms resulted in 860 data points. Higher agreement was found between pDES coordinates and CSD-based compared to DTI-based tractograms. ROC curves showed overall higher sensitivity at shorter distance cutoffs for CSD (8.5 mm) compared to DTI (14.5 mm). CSD-based CST tractograms showed significantly higher agreement with pDES, which was confirmed by linear modeling and posthoc tests ($P_{FWE} < 0.05$).

Conclusion

CSD-based CST tractograms were more accurate than DTI-based ones when validated using DES-based assessment of motor and sensory function. This demonstrates the potential benefits of structural mapping using CSD in clinical practice.

Clinical relevance statement

CSD-based tractograms of the CST are more sensitive than DTI-based tractograms when validated against sensory-motor DES mapping. This also demonstrated the feasibility of fully-automated CSD-based tractography for presurgical planning of the CST.

Introduction

Modern neurosurgery strives to optimize functional preservation and therapeutic outcomes. The gold-standard for function-preserving brain surgery[1] is intraoperative direct electrical stimulation (DES) during awake-surgery with image-guided neuronavigation. Magnetic resonance imaging (MRI)-based non-invasive brain mapping, normally complements and guides DES[2–5]. However, it may be the sole mapping method available in patients where the gold-standard is not feasible due to its complexity and potential complications[6–8].

Here we focused on presurgical mapping with diffusion MRI (dMRI)-based fiber tractography (FT). FT is the rendering of white matter fasciculi in 3D using dMRI signal contrast and computational modeling[9]. Due to its clinical accessibility and validation[10–14], the most commonly used FT approach in neurosurgical settings is diffusion tensor imaging (DTI)-based fiber assignment by continuous tracking (FACT)[15].

DTI-FACT suffers from multiple limitations making it suboptimal in neurosurgical practice. DTI assumes a single fiber direction per voxel, which compromises the accuracy of modeled streamline trajectories in voxels containing complex fiber architecture, leading to underestimation of the true extent of fiber bundles[16, 17]. DTI-FACT is also confounded by lesion-effects, as perilesional edema can compromise sensitivity to fiber orientation, resulting in false-negatives, i.e., missing streamlines/tracts. Additionally, FT suffers from user-bias and limited reproducibility when manual region of interest (ROIs) definition is used, which is the most common approach in clinical practice. These short-comings limit the accuracy of presurgical fiber tracking and motivate exploring alternative methods using automation, and higher-order model based tractography[16].

High angular-resolution imaging (HARDI) methods like constrained spherical deconvolution (CSD)[18] can address the limitations of DTI-FACT[19–25]. However, HARDI typically requires a more complex data acquisition compared to DTI[26–29]. Some techniques, such as CSD have been shown to improve tractography based on single-shell low angular-resolution clinical data[30, 31]. Probabilistic-FT tends to have a higher sensitivity compared to deterministic-FT, and thus may be preferable in a surgical setting where false-negatives are more critical. Anatomically-constrained

tractography (ACT)[32] may also improve accuracy of clinical tractography by constraining streamline origin/termination to the gray-white matter interface.

In this study we compared distance measures between DES coordinates acquired intraoperatively through pathology-tailored craniotomies and tractograms generated with deterministic DTI-FACT, probabilistic-DTI, and probabilistic-CSD, with and without ACT, to determine their respective suitability for presurgical mapping. We focused on the corticospinal tract (CST) and arcuate fasciculus (AF) given their relevance in clinical practice as primary pathways for sensory-motor and language functions.

Methodology

Research questions

We investigated the following research questions (RQ) in order to evaluate the performance of CSD- and DTI-based tractography with DES as the ground truth:

RQ1: How do distance measures between tractograms and DES coordinates compare between different tractography methods?

RQ2: How do these methods compare in binary agreement/disagreement with DES at different distance cutoffs?

RQ3: Are differences between tractography methods in the CST dependent on bundle volume and/or lesion volume?

Participants

We recruited 79 surgery-naïve patients referred for presurgical fMRI and DTI between 01/2019 and 01/2021, 22 patients also underwent intraoperative DES mapping (14 males, age=8 – 73 years, median=39.5, IQR=28, Neoplasms=18 and focal cortical dysplasia=4). Each patient was informed about the study and signed a written informed consent before participation, in accordance with the declaration of Helsinki. Local ethics committee approval was obtained (UZ/KU Leuven, Leuven, Belgium, S61759). Participating patients were excluded if they had undergone previous resective brain surgery, had brain implants, ventriculoperitoneal shunts, or had absolute contraindications to MRI. Detailed demographics and pathological information can be found in **S.table 1**.

MRI acquisition

Two 3-tesla MRI scanners were used for multimodal presurgical MRI scanning (Ingenia - Elition, and Achieva DStream, Philips Medical Systems, Best, The Netherlands), both with 32-channel receive head coils. The acquisition parameters for 3D T1-weighted images, T2- and T2 fluid attenuation inversion recovery (FLAIR) images were previously described [33]. Multi-shell dMRI data and reversed phase B0 images were acquired whenever feasible and tolerated by the patient, see **Table 1**. Single-shell data with or without reversed phase B0 images were used if multi-shell data couldn't be acquired, see **S.table 2**.

Table 1: MRI acquisition parameters			
	dMRI - part1	dMRI - part2	dMRI - part3
Pulse sequence	2D spin-echo EPI		
Acquisition plane	Axial		
TR/TE ms	4500/85		
FA °	85		
Voxel size in mm³	1.96*1.96*2.00		
Acquisition matrix	112*112*69		
In-plane acceleration	SENSE 1.6		
Multiband	3		
Pixel BW	2997		
PE direction	AP		
Fat shift direction	A		P
Bvalue (s/mm²)	1200	2500	0
Number of diffusion directions	127	124	0
Number of non-diffusion volumes	1	4 - 9	4 - 7
dMRI = diffusion magnetic resonance imaging, EPI = echo planar imaging, TR = repetition time, TE = echo time, ms = milliseconds, FA = flip angle, SENSE = SENSitivity encoding in plane parallel imaging acceleration, BW = bandwidth, PE = phase encoding, AP = anteroposterior, A = anterior, P = posterior			

Intraoperative brain mapping

Twenty patients underwent awake neurosurgery and DES, and 2 patients underwent DES with motor and somatosensory evoked potentials (MEP/SSEP). Intraoperative frameless neuronavigation (Curve, BrainLab, Munich, Germany) was employed in all cases. DES used the OSIRIS neurostimulator (Inomed Medizintechnik GmbH, Germany), and a bipolar fork stimulator with 5mm inter-electrode spacing for cortical mapping, and a monopolar suction-stimulator for subcortical mapping.

Data analysis

Figure 1 shows a schematic representation of the data preprocessing and analysis workflow. All acquired images were converted to the brain imaging data structure format (BIDS) [34] format using the KU Leuven Neuroimaging suite (KUL_NIS) [35] (https://github.com/treanus/KUL_NIS) and dcm2bids [36].

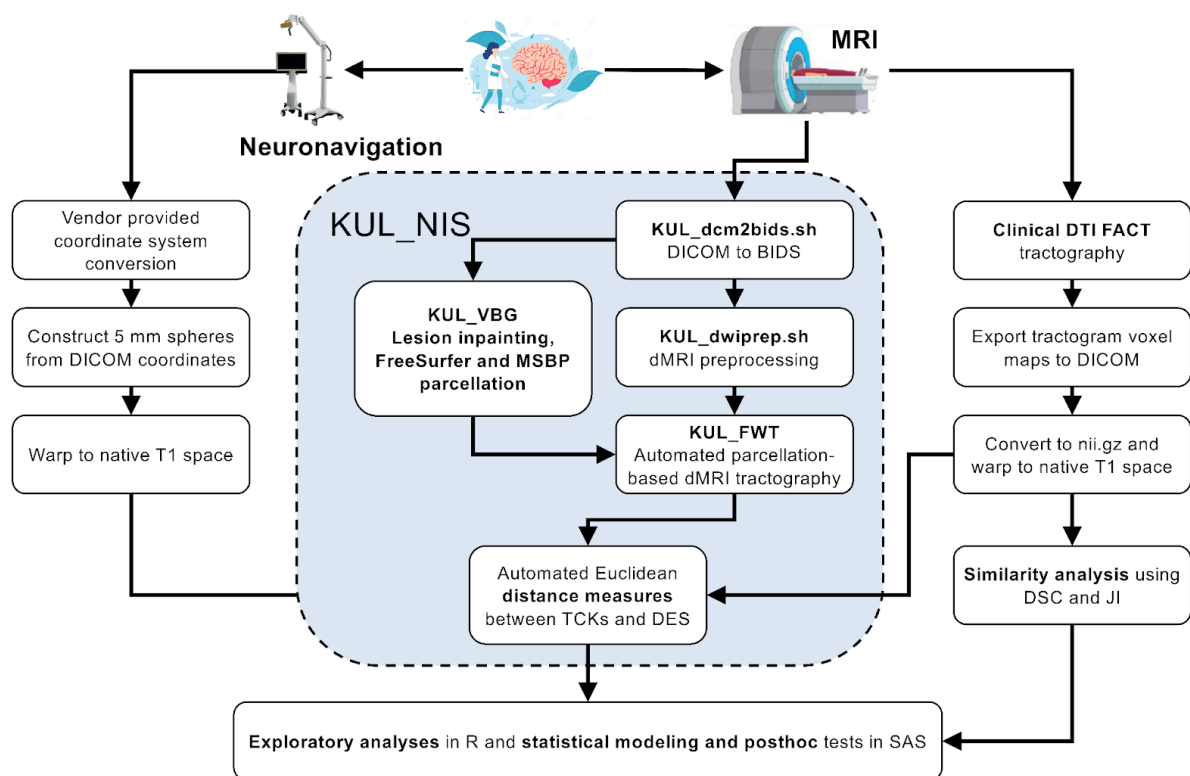


Figure 1: Schematic representation of the data preprocessing and analysis workflow used to compare different tractography results to intraoperative mapping outcome. MRI = magnetic resonance imaging, KUL_NIS = KU Leuven neuroimaging suite, BIDS = brain imaging data structure, KUL_VBG = KU Leuven virtual brain grafting, KUL_FWT = KU Leuven fun with tracts, DTI = diffusion tensor imaging, FACT = fiber assignment by continuous tracking, DSC = dice similarity coefficient, JI = Jaccard index

Structural images

Semi-automated classification in ITK-snap v3.8.0[37] was used for lesion segmentation using T1, T2, T2-FLAIR and contrast-enhanced T1-weighted images. KU Leuven Virtual brain grafting v0.52 (VBG)[33] (https://github.com/KUL-Radneuron/KUL_VBG) was used for T1 lesion-filling and T1 parcellation[38–40]. FreeSurfer[39] generated 1mm isotropic T1-weighted images were used as base images for calculating distances. Lesion volume, total intracranial volume (TIV), and lesion mask to TIV ratios were also calculated per patient.

Diffusion-weighted images

Routine clinical dMRI analysis involved in-line rigid interframe registration for motion correction on a clinical workstation (Philips Medical Systems, Best, The Netherlands). Advanced dMRI preprocessing used the KUL_dwiprep.sh script[35], which relies on FSL v6.0[41], ANTs v2.3.0[42, 43] and MRtrix3 v3.0.3[44] for denoising, correction of Gibbs ringing, subject motion, Eddy current artifacts, echo-planar imaging (EPI) distortion, and imaging bias. This was followed by DTI and CSD model fitting in MRtrix3[44–46]. FSL's topup[47] was used for EPI distortion correction if reversed phase B0 images were available, and synB0-DisCo v3.0[48] if not.

Tractography

Five different tractography methods were included: The standard clinical DTI-FACT based on manual white matter ROIs delineated[49] by 2 experienced operators on the Philips FiberTrack software was used as the reference method (FACT). Additionally, we included 4 automated approaches using probabilistic tensor-based tractography (TP)[50], and anatomically-filtered TP similar to ACT[32] (ATP), and probabilistic-CSD tractography using second order integration over fiber orientation distributions (iFOD2)[51], and anatomically-filtered iFOD2 (AiFOD2). This resulted in 5 different versions of every bundle of interest.

Automated tractography (TP, ATP, iFOD2, and AiFOD2) was done using the bundle-specific approach in KU Leuven Fun With Tracts v0.6 (KUL_FWT)[52] (https://github.com/KUL-Radneuron/KUL_FWT) with the addition of a model-based tractogram filtering step using streamline clustering (see supplementary methods). Automated tractography was done with 15,000(CST) and 6,000(AF) required number of streamlines for initial bundles. All generated tractograms were visually assessed for

quality before further analysis. ANTs[42, 43] registration was used to bring all results to each subject's T1 space for quantitative comparisons. Voxel-wise volumes for each bundle and bundle to TIV ratios were calculated per patient.

DES coordinates processing and distance measures

Saved DES coordinates were exported from the neuronavigator and used to create spheres with 5mm radii[53] in FSL[41] in the same space as the anatomical image used during surgery, then warped to T1 space with ANTs[42, 43] for comparison with the tractography results. Minimum Euclidean distances were calculated between all DES coordinates and all voxels of the corresponding tractogram in Python v3.8. Distance measures were rounded up to integers (in mm) because increments smaller than a single voxel (1mm) were not considered meaningful.

Statistical testing

Exploratory analysis

Distances between DES coordinates and tractograms can be considered akin to screening test results, and can thus be compared using confusion matrices, and ROC curves. However, due to the small sample size, and intersubject variation in number of DES coordinates, tractograms, and lesion size we opted for a more in-depth analysis. First (**RQ1**) we plotted the raw distance measures[54], then distances were averaged for each patient per DES response type to remedy the bias for patients with higher number of datapoints and ROC[55] curves were used to explore differences in sensitivity, specificity, and Youden's index[56] ideal cutoffs.

Binary agreement/disagreement rates (**RQ2**) at different distance cutoffs were calculated and assessed using confusion matrices. Tractogram-DES pairs with distance less than the cutoff represented positive datapoints, true-positives if involving pDES, and false-positives if nDES, and those with distance above the cutoff represented negative datapoints, true-negatives if involving nDES and false-negatives if involving pDES.

Two-part linear modeling

Distance data thresholded at the cutoff determined by ROC of all pooled data were used for a two-part linear mixed model and posthoc testing to compare FT methods while correcting for differences in relative volumes of tractograms and lesions per patient. The thresholded distances data were a semicontinuous variable with excess

zeros and an extremely right-skewed distribution, violating assumptions of normality. Therefore, and given the within-subjects nesting of repeated distance measures, we opted for a two-part model for longitudinal data[57, 58] (**RQ3**). The model was estimated with the %MIXCORR macro provided by Tooze, Gunwald and Jones, 2002[58] and PROC NLMIXED in SAS studio v9.4 (SAS Institute, Cary, NC, USA).

The first part (A) predicted the probability of overlap (distance=0), and the second part (B) predicted the distances between nonoverlapping (distance>0) tractogram-DES coordinate pairs. Distances were the dependent variable and DES response, and tractography method (FACT, TP, ATP, iFOD2, and AiFOD2) were used as predictors in both parts of the model. Bundle-to-TIV and lesion-to-TIV ratios, were used as covariates of noninterest. Adaptive Hochberg's[59] family-wise error-rate (FWE) correction was used to control for type(I) error in posthoc testing. Further detail can be found in supplementary material.

Results

Lesion segmentation, inpainting and tractography

Lesion masks had a median volume=44.70ml, minimum=1.20, maximum=232.129, and IQR=67.74ml. **Figure 2** shows the cumulative voxel-wise distribution of lesions in this sample of patients over the whole brain in MNI152 space. Lesion-inpainting and structural parcellation was successful in all patients. Tractography yielded 134 out of 135 attempted bundle reconstructions, with 1 failed right CST TP tractography (PT006). No bundles were excluded on inspection and none were repeated. **Figure 3 and S.figure 1** show all generated CST and AF tractograms, respectively. All methods showed a good level of visual agreement, but CSD-based methods resulted in more extensive tractograms. CSD captured the characteristic fanning of the CST into the lateral and inferior sensory-motor cortex, while DTI did not show this appearance, and AF CSD reconstructions reached further into the inferior frontal gyri and temporal lobes compared to DTI. As part of tractogram quality analysis we also explored the similarity of bundles generated with different FT methods, see supplementary material and **S.table 5** for details.

Intraoperative mapping and distance measures

DES mapping resulted in 51 pDES and 123 nDES coordinates, 3 patients had only nDES coordinates, namely PT011, PT012 and PT022, while 3 patients had only pDES coordinates, namely PT010, PT013, and PT015. **S.table 2** lists the DES tests, number of resulting pDES and nDES coordinates, elicited pDES response, acquired dMRI data, and bundle(s)-of-interest per patient. **Figure 4** shows images from 4 exemplar cases demonstrating the DES spheres and tractography results. Each tractogram was paired with pDES coordinates, and all nDES coordinates for distance measures, resulting in 860 distance measures with 174 measures for all methods except TP, which had 164 due to the failed right CST for PT006. **S.table 3** lists summarized descriptive statistics for distance measures, and **S.figure 2** shows the distribution of distances per bundle and DES response.

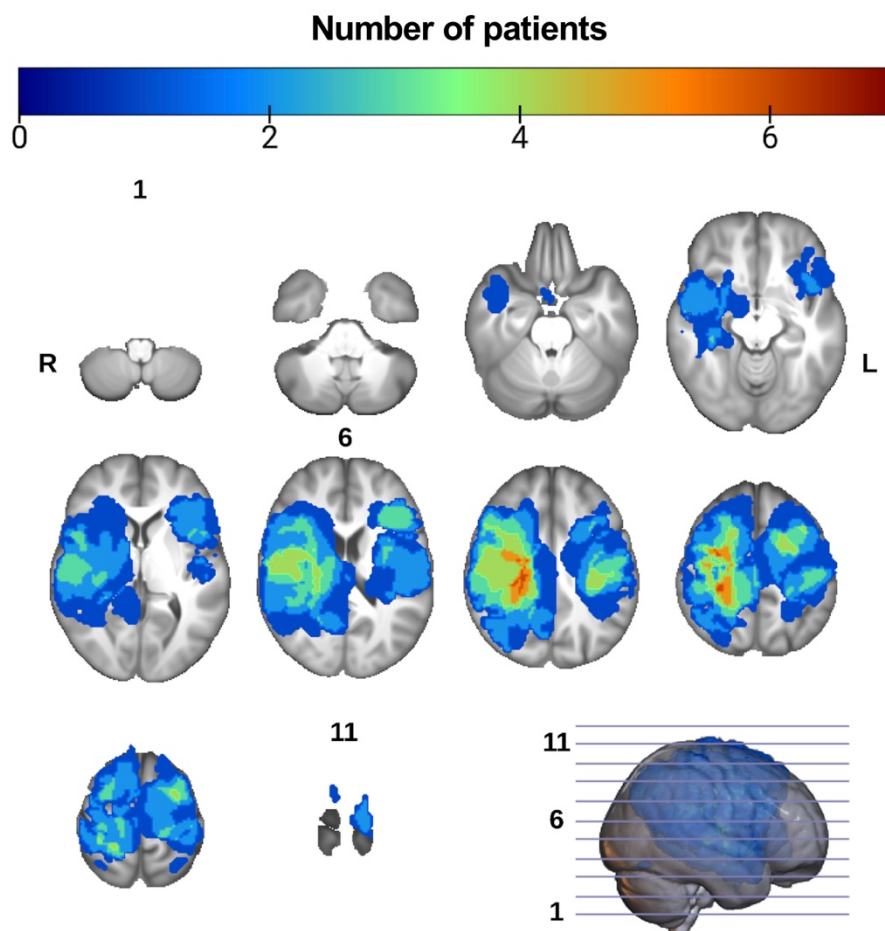


Figure 2: Spatial distribution of lesions from all patients overlaid onto the UK biobank T1 template brain in standard Montreal neurological institute (MNI) space. Perilesional edema was included in the masks of neoplasms if present. Overlay voxel intensities correspond to the sum of overlapping lesion masks from different patients. R = right, L = Left, slice numbers are indicated for the first, middle and last slices.

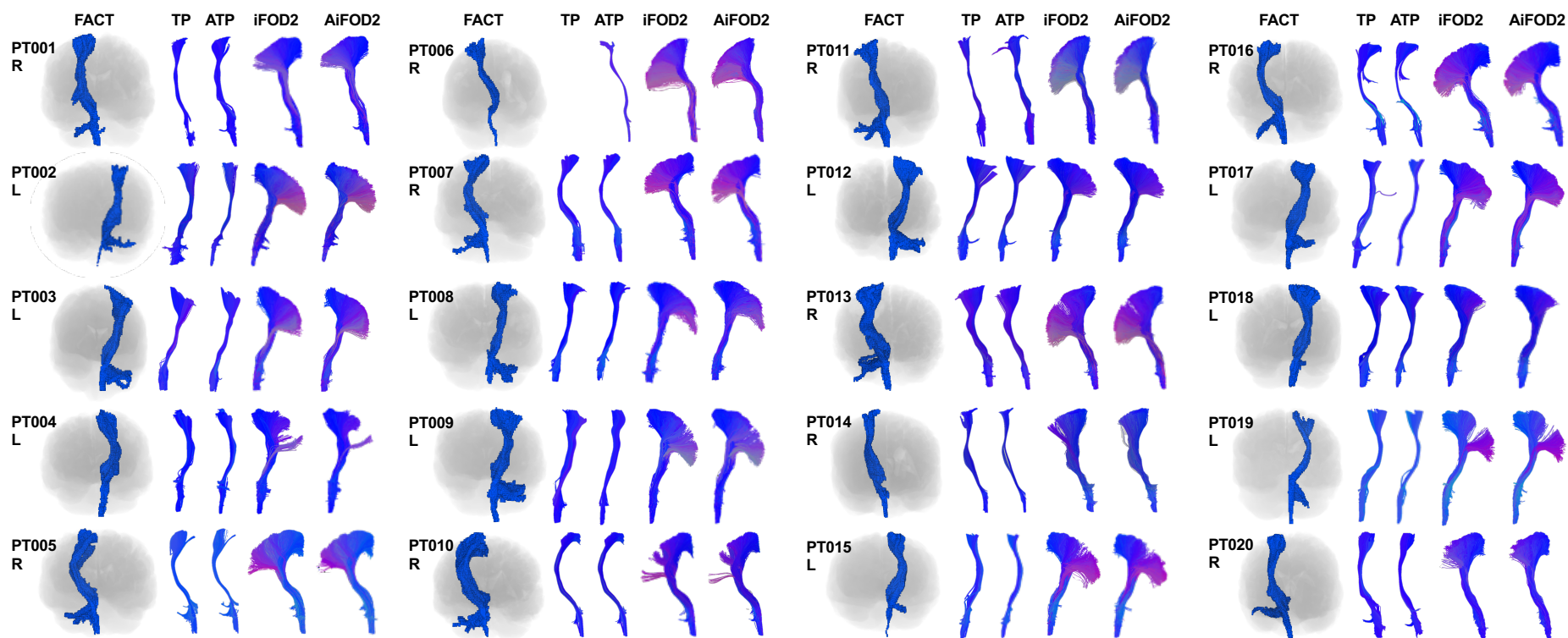


Figure 3: Corticospinal tractograms representative images **from** all methods in anterior view. The FACT tractogram outputs shown in blue are generated from volume rendered voxel masks with the T1-brain image silhouette shown underneath. The other 4 methods TP, ATP, iFOD2, and AiFOD2 are shown as 3D rendered streamlines with end-point directional color coding. All images are shown in radiological orientation. PT = patient, FACT = fiber assignment by continuous tracking, TP tensor probabilistic, ATP = anatomically constrained tensor probabilistic, iFOD2 = probabilistic tractography by second order integration over spherical harmonics, AiFOD2 = anatomically constrained iFOD2

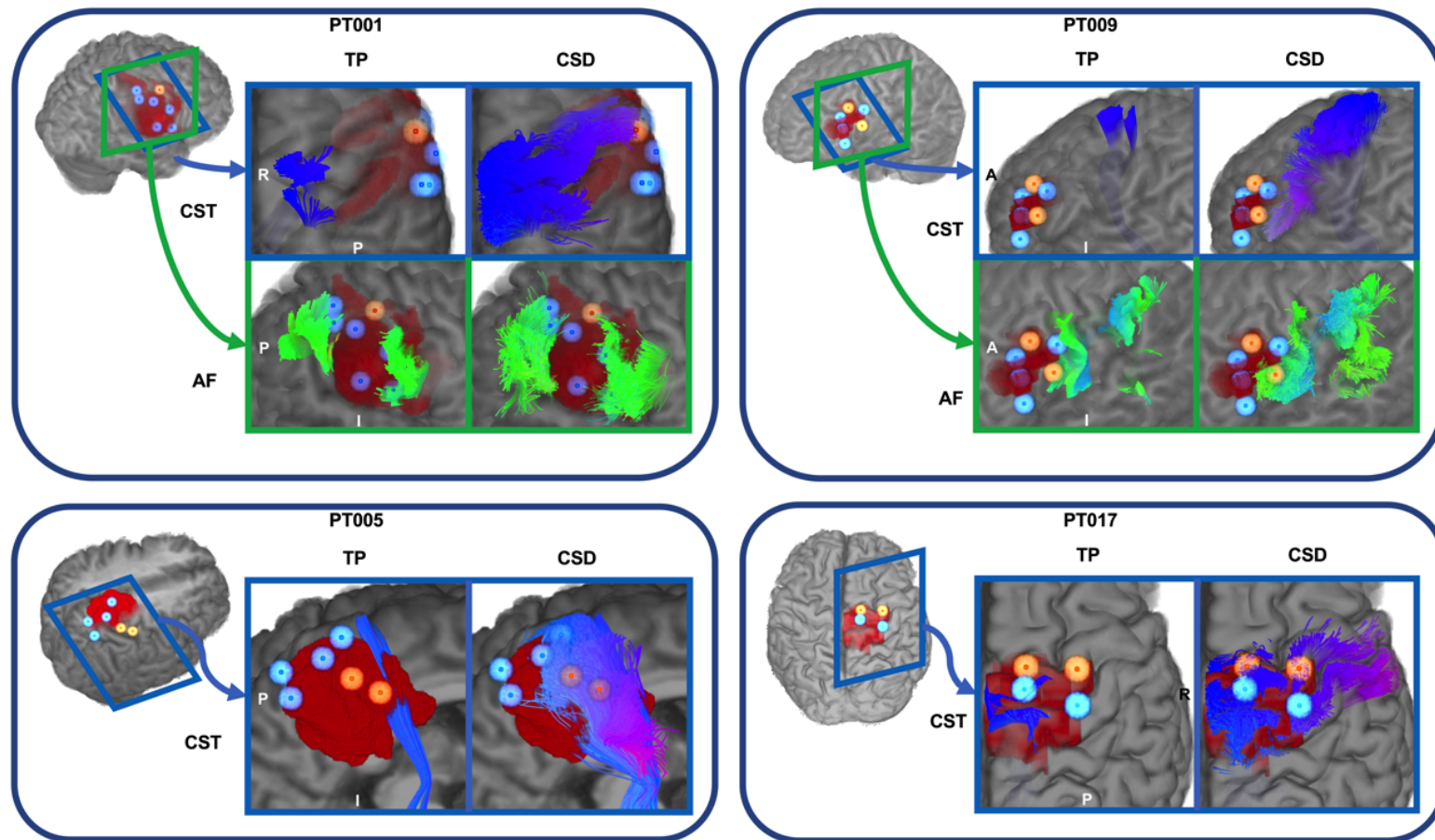


Figure 4: DES spheres, lesion masks and tractograms overlaid onto surface-rendered T1 images, for four example patients as generated with TP and iFOD2. Positive DES spheres are shown in semitransparent orange and their centers are shown in opaque orange, negative DES spheres are shown in semitransparent light blue with opaque blue centers. Tractograms are shown in endpoint-based directional color coding. TP = tensor probabilistic, iFOD2 = probabilistic tractography based on second order integration of spherical harmonic distributions, CST = corticospinal tract, AF = arcuate fasciculus, A = anterior, P = posterior, I = inferior, R = right

Statistical testing

Exploratory analysis

The CST and AF tractograms tended to be closer to pDES than to nDES coordinates, regardless of FT method. CSD-based tractograms showed shorter distances than DTI-based tractograms to both nDES and pDES coordinates (**RQ2**), see **S.figure 2** for plots of unthresholded distances. ROC curves for distance measures of all FT methods pooled together, as well as for both CSD-based and DTI-based FT methods are shown in **Figure 5**, and **S.figures 3** and **4** show the ROC curves for pooled raw distances, and each FT method separately with and without averaging.

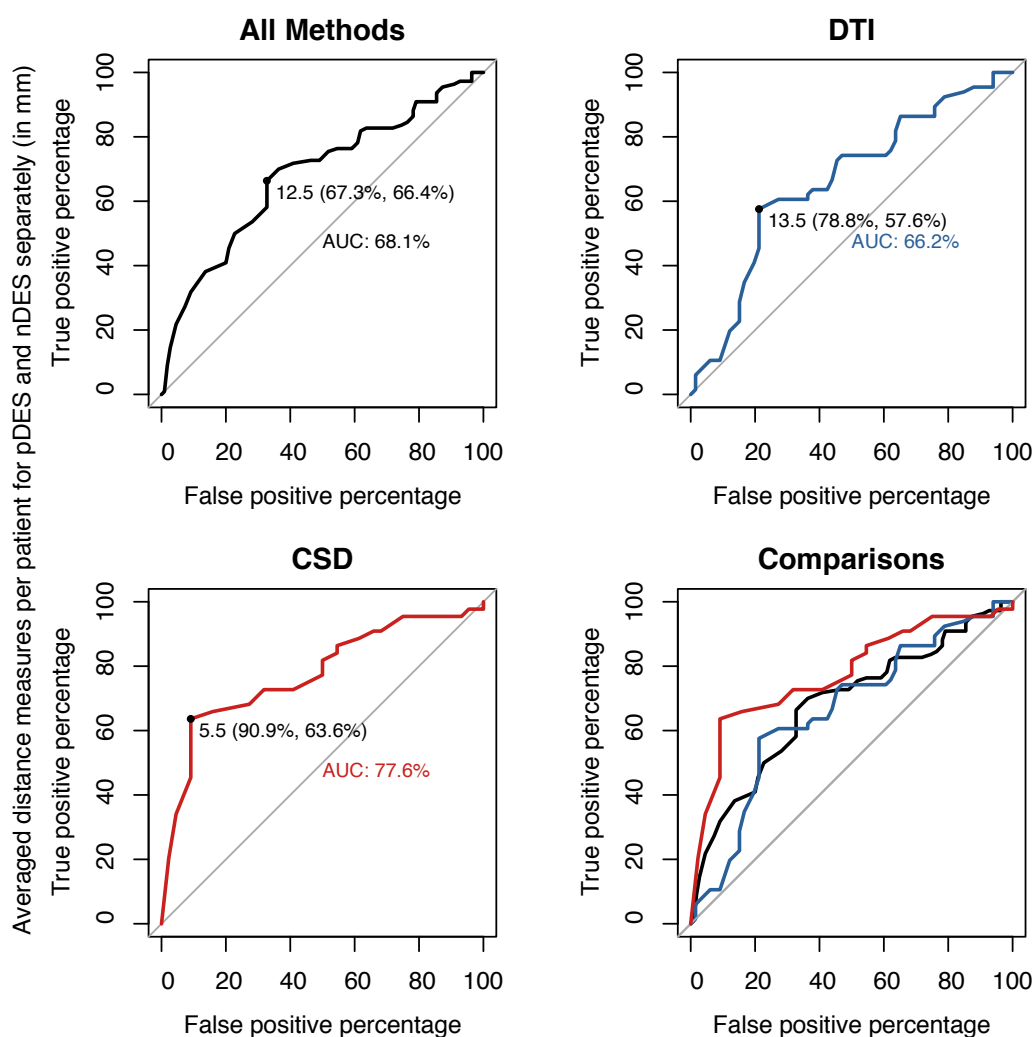


Figure 5: ROC curves and optimal distance cutoffs calculated using Youden's method. Results for pooled data from all methods are shown in black, DTI (FACT, TP and ATP) in blue, and CSD (iFOD2 and AiFOD2) in red, FACT = fiber assignment by continuous tracking, TP tensor probabilistic, ATP = anatomically constrained tensor probabilistic, iFOD2 = probabilistic tractography by second order integration over spherical harmonics, AiFOD2 = anatomically constrained iFOD2

All ROC curves showed notable increase in sensitivity and comparatively smaller decrease in specificity for CSD-based compared to DTI-based FT, but there were no significant differences in AUC on pairwise DeLong tests, see **S.table 4**. Youden's index determined distance cutoffs were also notably smaller for CSD-based methods (pooled=5.5mm, iFOD2=6.5mm, and AiFOD2=5.5mm) than DTI-based methods (pooled=13.5mm, FACT=19.5mm, TP=13mm and ATP=14.5mm), and the cutoff determined from all pooled data was 12.5mm. CSD-based methods were also associated with higher sensitivity at all ROC determined thresholds.

We used the 3 cutoffs determined from ROCs of pooled data of all methods, CSD, and DTI to define binary agreement/disagreement counts (**RQ2**) to compare FT methods using confusion matrices, see **Table 2**. These showed higher sensitivity and lower specificity for CSD-based compared to DTI-based tractography in both bundles at all distance cutoffs.

Table 2: Summarized binary agreement/disagreement confusion matrix results for all Tractogram-DES pairs						
FT methods	Thresholds (mm)	Accuracy	Sensitivity	Specificity	Pos Pred Value	Neg Pred Value
FACT	5.5	75%	33%	92%	63%	77%
	12.5	71%	61%	75%	50%	82%
	13.5	71%	65%	74%	51%	83%
TP	5.5	68%	39%	80%	45%	76%
	12.5	70%	67%	72%	49%	84%
	13.5	70%	67%	71%	49%	84%
ATP	5.5	69%	25%	87%	45%	74%
	12.5	72%	57%	78%	52%	81%
	13.5	72%	59%	77%	52%	82%
iFOD2	5.5	74%	65%	77%	54%	84%
	12.5	57%	88%	45%	40%	90%
	13.5	57%	90%	43%	40%	91%
AiFOD2	5.5	75%	71%	77%	56%	86%
	12.5	60%	92%	47%	42%	94%
	13.5	58%	92%	44%	41%	93%

FT = fiber tractography, Pos Pred = positive predictive, Neg Pred = negative predictive

Two-part model

The two-part linear model (**RQ3**) showed a significant main effect of tractography methods, and DES response type in both parts of the model. CSD-based methods (iFOD2 and AiFOD2) had larger probabilities of overlap (<12.5mm) and smaller distances (>12.5mm) compared to DTI-based methods (FACT, TP, ATP) for both pDES and nDES coordinates. Results also showed that pDES coordinates had significantly higher probability of overlap ($T(18) = -7.70$, $p < 0.001$), and were closer to tractograms than nDES coordinates were ($T(18) = -5.09$, $p < 0.001$). Bundle-to-TIV ratio had a significant effect ($T(18) = -2.55$, $p = 0.020$) only in the logistic part.

No significant interaction was found between DES response type and tractography method for the CST, indicating that differences between nDES and pDES were not significantly different between FT methods. Posthoc testing controlling for lesion-to-TIV and bundle-to-TIV ratios showed that CSD-based tractograms were more likely to overlap with and be closer to DES coordinates compared to DTI-based tractograms. These results are detailed in **Table 3** and **Figure 6**, and **S.figure 5** and **S.table 6** show results with 10.5mm distance cutoff.

Table 3: Results of post hoc tests after two-part linear modeling for the CST at 12.5 mm distance cutoff									
Test	Df	Model	T	P _{uncorr}	P _{FWE}	Model	T	P _{uncorr}	P _{FWE}
FACT v TP for DES and CST	18	A	-1.08	0.296	0.591	B	-1.27	0.220	0.783
FACT v ATP for DES and CST	18	A	-2.19	0.041	0.083	B	-2.36	0.030	0.195
FACT v iFOD2 DES and CST	18	A	5.98	<0.001	<0.001	B	5.51	<0.001	<0.001
FACT v AiFOD2 DES and CST	18	A	5.62	<0.001	<0.001	B	6.42	<0.001	<0.001
TP v ATP for DES and CST	18	A	-1.13	0.273	0.546	B	-1.01	0.324	0.324
TP v iFOD2 for DES and CST	18	A	6.75	<0.001	<0.001	B	6.49	<0.001	<0.001
TP v AiFOD2 for DES and CST	18	A	6.46	<0.001	<0.001	B	7.43	<0.001	<0.001
ATP v iFOD2 for DES and CST	18	A	7.59	<0.001	<0.001	B	7.51	<0.001	<0.001
ATP v AiFOD2 DES and CST	18	A	7.32	<0.001	<0.001	B	8.49	<0.001	<0.001
iFOD2 v AiFOD2 DES and CST	18	A	-0.44	0.665	0.665	B	-0.27	0.451	0.451

CST = corticospinal tract, FACT = fiber assignment by continuous tracking, TP = tensor probabilistic, ATP = anatomically constrained TP, iFOD2 = second order integration over fiber orientation distributions, AiFOD2 = anatomically constrained iFOD2, DF = degrees of freedom, T = t-statistic, P_{uncorr} = uncorrected p values, P_{FWE} = Hochberg family-wise error rate corrected p values

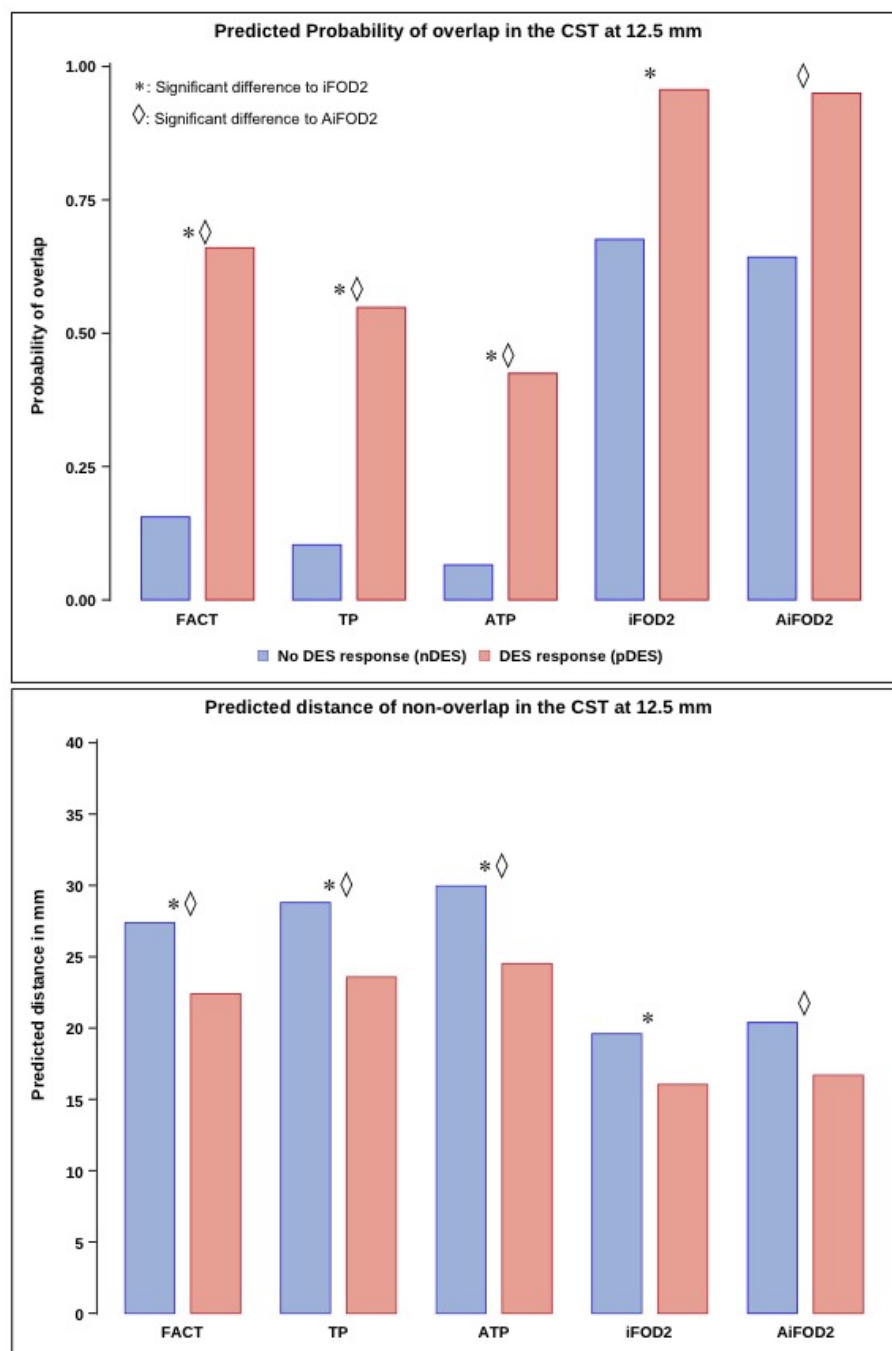


Figure 6: Bar plots for predicted probability of overlap between the CST and DES coordinates (Top) and predicted distances to nonoverlapping DES coordinates (Bottom) at 12.5 mm distance cutoff. CSD methods showed significantly higher probability of overlap, and lower distance if not overlapping compared to DTI methods. Differences between nDES and pDES were comparable between FT methods. CST = corticospinal tract, FACT = fiber assignment by continuous tracking, TP = tensor probabilistic, ATP = anatomically constrained TP, iFOD2 = probabilistic tractography by second order integration over spherical harmonics, AiFOD2 = anatomically constrained iFOD2

Discussion

The primary aim of this work was to compare the accuracy of CSD and DTI tractography results as evaluated against intraoperative DES stimulation in a group of 22 preoperative neurosurgical patients. First, raw distance measures from CSD-based tractograms to DES coordinates were smaller than for DTI tractograms. These also showed that pDES coordinates were generally closer to tractograms compared to nDES coordinates. ROC curve plots and Youden's index determined distance cutoffs, accuracy, sensitivity and specificity (**RQ1**) showed that pooled CSD methods (5.5mm) had a shorter distance cutoff compared to pooled DTI methods (13.5mm). This finding held with or without pooling and/or averaging. Overall higher sensitivity for CSD, and higher specificity for DTI tractograms was also found when comparing FT methods for their binary agreement/disagreement with DES at different distance cutoffs (**RQ2**) (5.5mm, 12.5mm, and 13.5mm).

The cutoff (12.5mm) determined by Youden's index on the averaged all pooled data ROC was used to define binary agreement/disagreement ($\leq 12.5 / > 12.5$ mm) rates between tractograms and DES. Lastly, we relied on a two-part linear mixed model and posthoc testing to investigate the effects of DES response, and differences between FT methods in probability of overlap (agreement) versus nonoverlap (disagreement), and distance measures for nonoverlapping tractogram-DES pairs while controlling for effects of variation in lesion and bundle volumes (**RQ3**).

This showed that regardless of FT method, pDES coordinates were significantly closer and more probable to overlap with the CST compared to nDES coordinates. In other words, brain locations giving positive sensory-motor functional responses on stimulation (pDES), so called eloquent areas, were more likely to overlap with the CST, and if not overlapping they would typically be closer than coordinates giving no functional effect on DES, so called non-eloquent. Furthermore, CST tractograms generated with CSD were significantly more probable to overlap with and be closer to DES coordinates compared to DTI, regardless of DES response. Posthoc testing confirmed that iFOD2 and AiFOD2 significantly outperformed FACT, TP and ATP in both parts of the model. These findings also indicate that the differences are not solely driven by larger volumes of CSD tractograms.

DTI-based tractography, when compared to DES, has previously been reported to have a generally satisfactory performance [11, 14, 60]. However, a closer look at the literature shows that these studies used manual tractography [61], and/or excluded negative DES coordinates from their analyses [62]. Additionally, the majority of previous work utilized comparatively simple statistical methods, and tended not to include covariates such as bundle volume, TIV or lesion volume in their analysis. In contrast to the literature, we have a less than satisfactory performance for DTI-based FT. This, most likely, is because we included nDES data in the analysis, which exposes the problem with DTI FT, especially on smaller distance cutoffs.

Few studies thus far have compared different tractography methods to DES and/or DBS results so far [63–66], and all found a clear benefit from using more advanced tractography methods such as CSD. Our study adds to this growing body of evidence showing that advanced FT can improve the presurgical representation of white matter anatomy based on dMRI. This study also showed that manual and automated DTI FT performed comparably, indicating that automated methods relying on accurate structural parcellation adapted to effects of pathology can be comparable to relying on an expert user for manual FT, but that both performed worse than CSD FT as far distances and agreement with DES coordinates was concerned.

The main limitation of this study is the small sample size, which motivated not including covariates for number of dMRI shells, pathology type, DES current, stimulator or approach used. However, the two-part linear model accounted for these factors by allowing for random intercepts and slopes per patient. Additionally, we did not include perioperative patient status in this analysis.

To summarize, both exploratory analyses and statistical modeling for the CST agreed that compared to DTI, CSD tractograms in this sample showed increased accuracy, with higher true-positives, at the cost of a smaller increase in false-positives, without significant reduction in true or false-negatives. No statistical testing was done for the AFs due to their small number (N=7) however it seemed to behave similar to the CST when evaluated qualitatively.

Bibliography

1. Gogos AJ, Young JS, Morshed RA, et al (2020) Awake glioma surgery: technical evolution and nuances. *J Neurooncol* 147:515–524. <https://doi.org/10.1007/s11060-020-03482-z>
2. Gharabaghi A, Tatagiba M (2007) [Expectations and demands of function preserving neurosurgery on imaging]. *Z Med Phys* 17:232–236. <https://doi.org/10.1016/j.zemedi.2007.09.006>
3. Dadario NB, Brahimaj B, Yeung J, Sughrue ME (2021) Reducing the Cognitive Footprint of Brain Tumor Surgery. *Front Neurol* 12:
4. Schebesch K-M, Rosengarth K, Brawanski A, et al (2019) Clinical Benefits of Combining Different Visualization Modalities in Neurosurgery. *Front Surg* 6:
5. Wengenroth M, Blatow M, Guenther J, et al (2011) Diagnostic benefits of presurgical fMRI in patients with brain tumours in the primary sensorimotor cortex. *Eur Radiol* 21:1517–1525. <https://doi.org/10.1007/s00330-011-2067-9>
6. Burnand C, Sebastian J (2014) Anaesthesia for awake craniotomy. *Contin Educ Anaesth Crit Care Pain* 14:6–11. <https://doi.org/10.1093/bjaceaccp/mkt024>
7. Dreier JD, Williams B, Mangar D, Camporesi EM (2009) Patients selection for awake neurosurgery. *HSR Proc Intensive Care Cardiovasc Anesth* 1:19–27
8. Jones H, Smith M (2004) Awake craniotomy. *Contin Educ Anaesth Crit Care Pain* 4:189–192. <https://doi.org/10.1093/bjaceaccp/mkh051>
9. Catani M, Howard RJ, Pajevic S, Jones DK (2002) Virtual in Vivo Interactive Dissection of White Matter Fasciculi in the Human Brain. *NeuroImage* 17:77–94. <https://doi.org/10.1006/nimg.2002.1136>
10. Javadi SA, Nabavi A, Giordano M, et al (2017) Evaluation of Diffusion Tensor Imaging–Based Tractography of the Corticospinal Tract: A Correlative Study With Intraoperative Magnetic Resonance Imaging and Direct Electrical Subcortical Stimulation. *Neurosurgery* 80:287–299. <https://doi.org/10.1227/NEU.0000000000001347>
11. Coenen VA, Jenkner C, Honey CR, Mädler B (2016) Electrophysiologic Validation of Diffusion Tensor Imaging Tractography during Deep Brain Stimulation Surgery. *Am J Neuroradiol* 37:1470–1478. <https://doi.org/10.3174/ajnr.A4753>
12. Nowacki A, Schlaier J, Debove I, Pollo C (2018) Validation of diffusion tensor imaging tractography to visualize the dentatorubrothalamic tract for surgical planning. *J Neurosurg* 130:99–108. <https://doi.org/10.3171/2017.9.JNS171321>
13. Küpper H, Groeschel S, Alber M, et al (2015) Comparison of different tractography algorithms and validation by intraoperative stimulation in a child with a brain tumor. *Neuropediatrics* 46:72–75. <https://doi.org/10.1055/s-0034-1395346>

14. Bello L, Castellano A, Fava E, et al (2011) Preoperative Diffusion Tensor Imaging (DTI): contribution to surgical planning and validation by intraoperative electrostimulation. In: Duffau H (ed) Brain Mapping. Springer Vienna, Vienna, pp 263–275
15. Mori S, Crain BJ, Chacko VP, van Zijl PC (1999) Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Ann Neurol* 45:265–269. [https://doi.org/10.1002/1531-8249\(199902\)45:2<265::aid-ana21>3.0.co;2-3](https://doi.org/10.1002/1531-8249(199902)45:2<265::aid-ana21>3.0.co;2-3)
16. Farquharson S, Tournier J-D, Calamante F, et al (2013) White matter fiber tractography: why we need to move beyond DTI. *J Neurosurg* 118:1367–1377. <https://doi.org/10.3171/2013.2.JNS121294>
17. Mormina E, Longo M, Arrigo A, et al (2015) MRI tractography of corticospinal tract and arcuate fasciculus in high-grade gliomas performed by constrained spherical deconvolution: Qualitative and quantitative analysis. *Am J Neuroradiol* 36:1853–1858. <https://doi.org/10.3174/ajnr.A4368>
18. Tournier J-DD, Calamante F, Connelly A (2007) Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* 35:1459–1472. <https://doi.org/10.1016/j.neuroimage.2007.02.016>
19. Wedeen VJ, Reese TG, Tuch DS, et al (2000) Mapping fiber orientation spectra in cerebral white matter with Fourier-transform diffusion MRI. In: Proceedings of the 8th Annual Meeting of ISMRM, Denver. p 82
20. Wedeen VJ, Hagmann P, Tseng W-YI, et al (2005) Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magn Reson Med* 54:1377–1386. <https://doi.org/10.1002/mrm.20642>
21. Wedeen VJ, Wang RP, Schmahmann JD, et al (2008) Diffusion spectrum magnetic resonance imaging (DSI) tractography of crossing fibers. *NeuroImage* 41:1267–1277. <https://doi.org/10.1016/j.neuroimage.2008.03.036>
22. Tuch DS, Reese TG, Wiegell MR, et al (2002) High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magn Reson Med* 48:577–582. <https://doi.org/10.1002/mrm.10268>
23. Tuch DS, Reese TG, Wiegell MR, Van J. Wedeen (2003) Diffusion MRI of Complex Neural Architecture. *Neuron* 40:885–895. [https://doi.org/10.1016/S0896-6273\(03\)00758-X](https://doi.org/10.1016/S0896-6273(03)00758-X)
24. Anderson AW (2005) Measurement of fiber orientation distributions using high angular resolution diffusion imaging. *Magn Reson Med* 54:1194–1206. <https://doi.org/10.1002/mrm.20667>
25. Tournier JD, Calamante F, Gadian DG, Connelly A (2004) Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage* 23:1176–1185. <https://doi.org/10.1016/j.neuroimage.2004.07.037>

26. Tournier J-D, Calamante F, Connelly A (2013) Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR Biomed* 26:1775–1786. <https://doi.org/10.1002/nbm.3017>
27. Dell'Acqua F, Tournier J-D (2019) Modelling white matter with spherical deconvolution: How and why? *NMR Biomed* 32:e3945. <https://doi.org/10.1002/nbm.3945>
28. Calamante F, Connelly A (2009) How many diffusion gradient directions are required for HARDI
29. Jeurissen B, Tournier J-DD, Dhollander T, et al (2014) Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* 103:411–426. <https://doi.org/10.1016/j.neuroimage.2014.07.061>
30. Toselli B, Tortora D, Severino M, et al (2017) Improvement in White Matter Tract Reconstruction with Constrained Spherical Deconvolution and Track Density Mapping in Low Angular Resolution Data: A Pediatric Study and Literature Review. *Front Pediatr* 5:182. <https://doi.org/10.3389/fped.2017.00182>
31. Calamuneri A, Arrigo A, Mormina E, et al (2018) White Matter Tissue Quantification at Low b-Values Within Constrained Spherical Deconvolution Framework. *Front Neurol* 9:716. <https://doi.org/10.3389/fneur.2018.00716>
32. Smith RE, Tournier J-D, Calamante F, Connelly A (2012) Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage* 62:1924–1938. <https://doi.org/10.1016/J.NEUROIMAGE.2012.06.005>
33. Radwan AM, Emsell L, Blommaert J, et al (2021) Virtual brain grafting: Enabling whole brain parcellation in the presence of large lesions. *NeuroImage* 229:117731. <https://doi.org/10.1016/j.neuroimage.2021.117731>
34. Gorgolewski KJ, Auer T, Calhoun VD, et al (2016) The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data* 3:1–9. <https://doi.org/10.1038/sdata.2016.44>
35. (2022) KULeuven Neuro Imaging Suite (KUL_NIS), https://github.com/treanus/KUL_NIS
36. Bedetti C, arnaudbore, Guay S, et al (2022) UNFmontreal/Dcm2Bids: 2.1.7, <https://github.com/UNFmontreal/Dcm2Bids>
37. Yushkevich PA, Pashchinskiy A, Oguz I, et al (2019) User-Guided Segmentation of Multi-modality Medical Imaging Datasets with ITK-SNAP. *Neuroinformatics* 17:83–102. <https://doi.org/10.1007/s12021-018-9385-x>
38. Henschel L, Conjeti S, Estrada S, et al (2020) FastSurfer - A fast and accurate deep learning based neuroimaging pipeline. *NeuroImage* 219:117012–117012. <https://doi.org/10.1016/j.neuroimage.2020.117012>

39. Fischl B (2012) FreeSurfer. *NeuroImage* 62:774–781. <https://doi.org/10.1016/j.neuroimage.2012.01.021>
40. Tourbier S, Aleman-Gomez Y, Griffa A, et al (2019) *sebastientourbier/multiscalebrainparcellator: Multi-Scale Brain Parcellator v1.1.1*
41. Jenkinson M, Beckmann CF, Behrens TEJ, et al (2012) FSL. *NeuroImage* 62:782–790. <https://doi.org/10.1016/j.neuroimage.2011.09.015>
42. Avants BB, Tustison NJ, Song G, et al (2011) A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *Neuroimage* 54:2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025.A>
43. Tustison NJ, Cook PA, Holbrook AJ, et al (2021) The ANTsX ecosystem for quantitative biological and medical imaging. *Sci Rep* 11:9068. <https://doi.org/10.1038/s41598-021-87564-6>
44. Tournier J-D, Smith RE, Raffelt DA, et al (2019) MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage* 202:116137–116137. <https://doi.org/10.1016/j.neuroimage.2019.116137>
45. Dhollander T, Raffelt D, Connelly A (2016) Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image
46. Dhollander T, Raffelt D, Connelly A (2018) Accuracy of response function estimation algorithms for 3-tissue spherical deconvolution of diverse quality diffusion MRI data
47. Andersson JLR, Skare S, Ashburner J (2003) How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20:870–888. [https://doi.org/10.1016/S1053-8119\(03\)00336-7](https://doi.org/10.1016/S1053-8119(03)00336-7)
48. Schilling KG, Blaber J, Huo Y, et al (2019) Synthesized b0 for diffusion distortion correction (Synb0-DisCo). *Magn Reson Imaging* 64:62–70. <https://doi.org/10.1016/j.mri.2019.05.008>
49. Catani M, Thiebaut de Schotten M (2008) A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex* 44:1105–1132. <https://doi.org/10.1016/j.cortex.2008.05.004>
50. Jones DK (2008) Tractography Gone Wild: Probabilistic Fibre Tracking Using the Wild Bootstrap With Diffusion Tensor MRI. *IEEE Trans Med Imaging* 27:1268–1274. <https://doi.org/10.1109/TMI.2008.922191>
51. Tournier J-D, Calamante F, Connelly A (2010) Improved probabilistic streamlines tractography by 2nd order integration over fibre orientation distributions. *Proc Intl Soc Mag Reson Med ISMRM* 18:
52. Radwan AM, Sunaert S, Schilling K, et al (2022) An atlas of white matter anatomy, its variability, and reproducibility based on constrained spherical deconvolution of

- diffusion MRI. NeuroImage 254:119029.
<https://doi.org/10.1016/j.neuroimage.2022.119029>
53. Cochereau J, Deverdun J, Herbet G, et al (2016) Comparison between resting state fMRI networks and responsive cortical stimulations in glioma patients. *Hum Brain Mapp* 37:3721–3732. <https://doi.org/10.1002/hbm.23270>
 54. Patil I (2021) Visualizations with statistical details: The “ggstatsplot” approach. *J Open Source Softw* 6:3167. <https://doi.org/10.21105/joss.03167>
 55. Robin X, Turck N, Hainard A, et al (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. <https://doi.org/10.1186/1471-2105-12-77>
 56. Youden WJ (1950) Index for rating diagnostic tests. *Cancer* 3:32–35. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)
 57. Farewell VT, Long DL, Tom BDM, et al (2017) Two-Part and Related Regression Models for Longitudinal Data. *Annu Rev Stat Its Appl* 4:283–315. <https://doi.org/10.1146/annurev-statistics-060116-054131>
 58. Tooze JA, Grunwald GK, Jones RH (2002) Analysis of repeated measures data with clumping at zero. *Stat Methods Med Res* 11:341–355. <https://doi.org/10.1191/0962280202sm291ra>
 59. Hochberg Y, Benjamini Y (1990) More powerful procedures for multiple significance testing. *Stat Med* 9:811–818. <https://doi.org/10.1002/sim.4780090710>
 60. Manan AA, Yahya N, Idris Z, Manan HA (2022) The Utilization of Diffusion Tensor Imaging as an Image-Guided Tool in Brain Tumor Resection Surgery: A Systematic Review. *Cancers* 14:2466. <https://doi.org/10.3390/cancers14102466>
 61. Xiao X, Kong L, Pan C, et al (2021) The role of diffusion tensor imaging and tractography in the surgical management of brainstem gliomas. *Neurosurg Focus* 50:E10. <https://doi.org/10.3171/2020.10.FOCUS20166>
 62. Berman JI, Berger MS, Chung SW, et al (2007) Accuracy of diffusion tensor magnetic resonance imaging tractography assessed using intraoperative subcortical stimulation mapping and magnetic source imaging. *J Neurosurg* 107:488–494. <https://doi.org/10.3171/JNS-07/09/0488>
 63. Sheng Z, Yu J, Chen Z, et al (2021) Constrained-Spherical Deconvolution Tractography in the Evaluation of the Corticospinal Tract in Glioma Surgery. *Front Surg* 8:646465. <https://doi.org/10.3389/fsurg.2021.646465>
 64. Gröschel S, Kumar V, Schuhmann MU, et al (2011) Comparison of different tractography algorithms and validation by intraoperative stimulation in a child with a brain tumor. *Neuropediatrics* 42:P044. <https://doi.org/10.1055/s-0031-1274016>

65. Bucci M, Mandelli ML, Berman JI, et al (2013) Quantifying diffusion MRI tractography of the corticospinal tract in brain tumors with deterministic and probabilistic methods. *NeuroImage Clin* 3:361–368. <https://doi.org/10.1016/j.nicl.2013.08.008>
66. Mandelli ML, Berger MS, Bucci M, et al (2014) Quantifying accuracy and precision of diffusion MR tractography of the corticospinal tract in brain tumors: Clinical article. *J Neurosurg* 121:349–358. <https://doi.org/10.3171/2014.4.JNS131160>