Automated detection of axonal injury along white matter tracts in acute severe head trauma

Chiara Maffei[1,2], Natalie Gilmore[2], Samuel B. Snider[3], Yelena G. Bodien[2,4], Anastasia Yendiki[1], Brian L. Edlow[1,2]

[1] Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, MA, USA
[2] Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston, MA, USA
[3] Department of Neurology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[4] Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Harvard Medical School, Charlestown, MA, USA.

Corresponding author:
Chiara Maffei, Ph.D.
Athinoula A. Martinos Center for Biomedical Imaging
Massachusetts General Hospital and Harvard Medical School
149 13th Street, Charlestown
Boston, MA, 02129

Key Words:
TBI; Tractography; Traumatic Axonal Injury; Diffusion MRI; White Matter

Abstract

New techniques for individualized assessment of white matter integrity are needed to detect traumatic axonal injury (TAI) and predict outcomes in critically ill patients with acute severe traumatic brain injury (TBI). Diffusion MRI tractography has the potential to quantify white matter microstructure *in vivo* and has been used to characterize tract-specific changes in TBI. However, tractography is not routinely used in the clinical setting to assess the extent of TAI, in part because focal lesions reduce the robustness of automated methods. In the present study, we propose an automated pipeline for individualized assessment of white matter damage in patients with acute severe TBI that is based on tractography reconstructions and multivariate analysis of along-tract diffusion metrics. We used the Mahalanobis distance to test for a deviation of diffusion metrics along 40 white matter tracts in 18 patients with acute severe TBI as compared to 33 healthy subjects. The automated pipeline successfully reconstructed white matter tracts in all patients with a successful FreeSurfer anatomical segmentation (17 of 18 patients), including 13 with focal lesions. In these 17 patients, a mean of 37.5 +/- 2.1 tracts were reconstructed without the need for manual intervention and a mean of 2.47 +/- 2.1 needed to be reinitialized upon visual inspection. The pipeline detected at least one injured tract in all patients (mean: 12.6, SD: 7.6). The number and neuroanatomic location of disrupted tracts varied across patients and levels of consciousness. The pre-motor, parietal, and temporal sections of the corpus callosum were the structures most frequently injured (in 10, 9, and 8 patients respectively), consistent with histological studies of TAI. Multivariate measures of TAI did not show a significant association with behavioral measures of consciousness, or subscales measuring basic language and motor function. In summary, we provide proof-of-principle evidence that an automated tractography pipeline can be used to detect and quantify TAI in patients with acute severe head trauma and could therefore assist in the clinical assessment of TAI.

1. Introduction

Traumatic axonal injury (TAI) is the most common pathologic substrate of head trauma (Brody et al., 2015; McGinn & Povlishock, 2016) and has a strong association with adverse clinical outcomes (Benson et al., 2007; Gennarelli et al., 1982; Moe et al., 2020; Newcombe et al., 2010). TAI is caused by high-velocity translational and rotational forces that stretch and shear axons, leading to either primary axotomy or a secondary cascade that can result in axonal degeneration (Hill et al., 2016). TAI lesions are typically multifocal, with a neuroanatomic distribution that varies based on the mechanism of injury (Adams et al., 1989; Blumbergs et al., 1995; Johnson et al., 2013). For example, individuals in high-speed motor vehicle accidents are more likely to experience TAI in the deep subcortical white matter than are people who are assaulted, whose burden of TAI is more likely to be limited to the hemispheric white matter (Adams et al., 1989; Johnson et al., 2013). Neuroimaging studies that compare patients at the group level are insensitive to these variations in lesion location and therefore are suboptimal for assessing individual patients in clinical care (Betz et al., 2012). For clinicians and the families of patients with acute severe traumatic brain injury (TBI), the inability to measure the burden of TAI creates prognostic uncertainty and complicates discussions about continuation of life-sustaining therapy in the intensive care unit (ICU) (Izzy et al., 2013; Turgeon et al., 2011). New techniques for individualized assessment of white matter (WM) injury (Jolly et al., 2020; Kim et al., 2013; Yuh et al., 2014) are therefore needed to enhance detection of TAI and improve the accuracy of outcome prediction in critically ill patients with acute severe TBI.

Due to its diffuse and microscopic nature (Adams et al., 1989; Blumbergs et al., 1995; Johnson et al., 2013), TAI is challenging to detect and quantify using routine imaging modalities, such as CT and conventional MRI (Betz et al., 2012). Diffusion MRI (dMRI) possesses the unique ability to probe WM tissue microstructure non-invasively (Basser et al., 1994), making it possible to identify TAI *in vivo* (Hashim et al., 2017; Li et al., 2011; Xu et al., 2007). Numerous studies have used measures obtained from the diffusion tensor (DT) model to assess WM integrity in patients with TBI (Arfanakis et al., 2002; Asken et al., 2017; Galanaud et al., 2012; Hulkower et al., 2013; Inglese et al., 2005; Newcombe et al., 2007; Zhang et al., 2017) and have shown its sensitivity to detect WM abnormalities across severities (*i.e.,* mild, moderate, severe) and recovery phases (*i.e.,* acute, subacute, chronic). Typically, summary DT metrics such as fractional anisotropy (FA) and mean diffusivity (MD) are extracted from WM regions of interest (ROIs) that are manually drawn in subject's space (Inglese et al., 2005; Mac Donald et al., 2011) or segmented from an atlas (Galanaud et al., 2012; O'Phelan et al., 2018). Whole-brain voxel-based analysis (VBA) (Van Hecke et al., 2016) and tract-based spatial statistics (Smith et al., 2006) approaches are also used to compare

normalized DT scalar maps between patient groups (Kinnunen et al., 2011), or between a control group and a patient (Jolly et al., 2020). While these approaches provide flexibility to extract DT metrics from WM regions and have demonstrated sensitivity to TAI (Jolly et al., 2020; Perlbarg et al., 2009), they do not localize injury along the trajectory of white matter tracts at the individual level. Moreover, these methods rely on accurate registration between patients and an atlas, which can be challenging in the presence of lesions and may not reflect interindividual anatomical differences. They also may involve manual drawing of ROIs, which can be time-consuming and infeasible in clinical contexts.

Diffusion-based tractography allows for the delineation and visualization of the 3D profiles of WM pathways in subject space. Tractography methods provide measurements of the magnitude and distribution of axonal damage along specific WM tracts, and thus may have additional prognostic utility for guiding early decisions about goals of care and continuation of life-sustaining therapy. Indeed, a patient's potential to recover consciousness and cognitive function depends upon the integrity of brain networks probed by tractography (Edlow et al., 2021). However, few studies have utilized tractography in patients with acute TBI (D'Souza et al., 2015; Ordóñez-Rubiano et al., 2017; Snider et al., 2019; Wang et al., 2008; Warner et al., 2010) and even fewer have focused on critically ill patients with acute severe TBI (Ordóñez-Rubiano et al., 2017; Snider et al., 2019; Wang et al., 2008). dMRI tractography techniques are not routinely used in the clinical setting to assess the extent of TAI (Schweitzer et al., 2019) because the presence of focal lesions is often a challenge for these techniques.

We aimed to test a robust pipeline for individualized TAI assessment in patients with acute severe TBI that combines automated reconstruction of WM tracts in subject space with a multivariate analysis of along-tract DT metrics. Several considerations are important in this application. Manual delineation of ROIs for tractography is time-intensive and operator-dependent (Rheault et al., 2020; Schilling et al., 2021). Automated tractography methods that use ROIs from an atlas rely on accurate registration between the patient and the atlas, which can be difficult to achieve in the presence of focal lesions (Crinion et al., 2007; Ripollés et al., 2012). Along-tract measures can provide higher sensitivity than diffusion metrics averaged over the entire extent of the tract (Colby et al., 2012; Pieri et al., 2021). Finally, local tractography methods, which reconstruct a tract step by step, by considering only the local diffusion orientation at each step, can be confounded by lesions and unable to step through them.

We used TRACULA (TRActs Constrained by UnderLying Anatomy), a method for global probabilistic tractography with anatomical neighborhood priors (Yendiki et al., 2011), which allowed us to automatically reconstruct 40 WM tracts at the individual level (Maffei et al., 2021). The anatomical neighborhood priors in TRACULA encode information about the relative position of the tracts with respect

to their surrounding anatomical structures, rather than their absolute coordinates in an atlas space. Therefore, TRACULA does not necessitate accurate registration to an atlas (Zollei et al, 2019; Maffei et al., 2021). It uses global tractography (*i.e.,* it represents the tract as a spline and fits its shape to the diffusion orientations from the entire brain). Thus, TRACULA would not "stop" at a lesion the way that a local tractography method might. We previously showed that TRACULA, when trained on very high-quality data from the Human Connectome Project (HCP) (Fan et al., 2015), improves the accuracy of tractography reconstructions in routine quality data when compared to a common multi-ROI tractography approach (Maffei et al., 2021). These methodologic advantages of TRACULA suggest potential for translation of this technique to critically ill patients with acute severe TBI, who are not medically stable enough to be scanned with a time-intensive, high-quality dMRI sequence.

We applied TRACULA to a diffusion MRI dataset prospectively acquired in a cohort of 18 critically ill patients with acute severe TBI who were imaged on a clinical 3 Tesla MRI scanner. We used TRACULA to identify the extent and location of injured WM tracts in each patient compared to 33 healthy subjects. We focused on along-tract measures of FA and MD, as these microstructural measures are the most commonly studied in the field of TBI (Hulkower et al., 2013). The goals of this study were: i) to demonstrate the feasibility of applying the TRACULA automated tractography stream (Maffei et al., 2021; Yendiki et al., 2011) in patients with acute severe TBI, including those with focal lesions, using dMRI data acquired on a clinical scanner; ii) to implement an individualized multivariate approach to measure patient-specific TAI severity, using reconstructions derived from the automated TRACULA tractography pipeline; iii) to test for associations between bundle-specific dMRI measures of TAI and behavioral measures of consciousness; and iii) to explore the relationship between TAI within domain-specific tracts and subscales measuring basic language and motor function.

2. Methods

2.1 Participants

We prospectively enrolled 18 patients (mean age: 28.6 +/- 8.7 years, 13 male) with acute severe TBI, sixteen of whom were enrolled as part of a previously described pilot study (Edlow et al., 2017), and two of whom were subsequently enrolled as pilot cases for an ongoing observational study (ClinicalTrials.gov NCT03504709) (See Table 2 for additional information). Inclusion criteria were: 1) age 18 to 65 years; and 2) traumatic coma, defined by Glasgow Coma Scale (GCS) (Teasdale & Jennett, 1974) total score of six without eye opening on at least one neurologic examination before ICU admission; and 3) no eye opening for 24 hours after injury. Exclusion criteria were: 1) prior history of severe brain injury

or neurodegenerative disease; 2) life expectancy less than six months per physician judgment; 3) presence of metal contraindicating MRI; and/or 4) no pre-injury English fluency.

We also enrolled 33 sex- and age-matched healthy control subjects (mean age 31.2 years, 21 male) with no history of neurological, psychiatric, cardiovascular, pulmonary, renal or endocrinological disease (Edlow et al., 2017). We confirmed that age was not statistically different between the two groups using a nonparametric Wilcoxon sum rank test in Python 3 using SciPy Stats 1.3.1 (Jones et al., 2001) ($W = -1.116$, $p = 0.2643$). Written informed consent was obtained from healthy subjects and from patients' surrogate decision-makers in accordance with a study protocol approved by the Mass General Brigham Institutional Review Board.

2.2 Standardized Assessment

Immediately prior to MRI, each patient's level of consciousness was prospectively assessed by a neurologist on the investigator team (B.L.E.) via behavioral evaluation with the GCS (Teasdale & Jennett, 1974) and the Coma Recovery Scale-Revised (CRS-R) (Giacino et al., 2004). Per routine ICU care at our institution, neurological examinations and GCS assessments were performed off of sedation every two-to-four hours (depending on clinical stability and safety considerations) by treating clinicians. ICU clinician GCS assessments were used to determine time from injury to coma emergence (*i.e.*, eye opening or localization to noxious stimulation) and time to command-following. If a patient's neurological examination fluctuated, the first behavioral evidence of coma emergence and command-following were reported.

2.3 MRI Data Acquisition

MRI data were acquired as soon as the treating clinicians deemed patients stable for transport. Diffusion weighted imaging data were acquired in the ICU on a 3T Skyra scanner (Siemens Medical Solutions, Malvern, PA) with a 32-channel head coil. Three healthy subjects and eight patients were scanned using an echo-planar imaging (EPI) sequence with the following parameters: $2x2x2\ mm$, $60\ b = 2{,}000\ s\ /\ mm^2$ and $10\ b = 0\ s\ /\ mm^2$ volumes, $TR = 13{,}700$ ms, and $TE = 98\ ms$. Thirty healthy subjects and 8 patients were scanned using an EPI sequence with simultaneous multi-slice (SMS) acceleration (Setsompop et al., 2012) and the following parameters: $2x2x2\ mm$, $60\ b = 2000\ s/mm^2$ and $10\ b = 0\ s\ /\ mm^2$, $TR = 6{,}700\ ms$, $TE = 100\ ms$, acceleration factor = 2. Previous work from our group showed that diffusion-derived connectivity metrics did not differ significantly between these two acquisitions (Snider et al., 2019). The length of the diffusion sequence was 8 minutes and 46 seconds. Multi-echo MPRAGE structural images were also acquired with the following parameters: $1x1x1\ mm$,

acquisition matrix = 256×256, $TE = 1.69\,ms$, $TR\ 2530\,ms$, $flip\ angle = 7°$. All patients and healthy subjects were imaged using the same scanner and head coil.

2.4 Diffusion MRI Data Processing

Diffusion weighted images were skull stripped and corrected for eddy current distortions and movement in FSL 6.0.1 (Andersson & Sotiropoulos, 2015). The tensor and the ball and stick model were fit to the data using DTIFIT and BEDPOSTX (Behrens et al., 2003) in FSL 6.0.1, respectively. Automated reconstruction of 40 major WM pathways was performed using the global probabilistic tractography algorithm TRACULA in FreeSurfer 7.2.0 (Maffei et al., 2021; Yendiki et al., 2011). Although TRACULA can reconstruct a total of 42 WM tracts, we decided to exclude the right and left fornix, because optimal identification of these WM bundles requires segmentation of thalamic subnuclei that are not provided by the standard "recon-all" FreeSurfer processing pipeline. The mathematical formulation of TRACULA has been described elsewhere (Yendiki et al., 2011, 2016). Briefly, the algorithm models a pathway as a cubic spline, which is initialized with the median streamline of the training set. A random sampling algorithm is used to draw samples from the posterior probability distribution of the pathway, which is decomposed into the likelihood term and the prior term. The likelihood term fits the shape of the spline to the diffusion orientations obtained from the ball-and-stick model in the voxels that it goes through. The prior term fits the shape of the spline to its anatomical neighborhood, given the manually labeled examples of this pathway from the training set and the anatomical segmentation volumes of both test and training subjects. The number of control points of the cubic spline were chosen as previously described (Maffei et al., 2021). Along-tract tensor-derived metrics obtained from DTIFIT were extracted for each of the reconstructed tracts using a pointwise assessment of streamline tractography attributes (PASTA) (Jones et al., 2005). For each of the 40 tracts, 1D along-tract profiles of FA and MD were generated by projecting the value of each measure from every point on every automatically reconstructed streamline to its nearest point on a reference streamline, as previously described (Maffei et al., 2021). The reference streamline is the mean of the training streamlines for each tract, ensuring that all subject data were sampled at the same number of cross-sections along a given bundle. Before extracting the diffusion measures, the posterior probability distribution estimated by TRACULA was thresholded by masking out all values below 20% of the maximum, which is the default threshold in TRACULA (Maffei et al., 2021; Yendiki et al., 2011; Zöllei et al., 2019). All the pre-processing steps detailed in this section are performed automatically by the TRACULA pipeline.

2.5 Structural MRI Data Processing

The anatomical segmentation that was needed to compute the anatomical neighborhood priors in TRACULA was obtained by analyzing the structural T1-weighted image in FreeSurfer 7.1.1 (Dale et al., 1999; Fischl, 2012). The automated "recon-all" pipeline was run with default parameters, except for the "bigventricle" option, which was added to optimize segmentation in patients that may have enlarged ventricles post-injury. The obtained segmentation volumes included a combination of the Desikan-Killiany cortical parcellation labels (Desikan et al., 2006) and the standard FreeSurfer subcortical segmentation. To compute the prior probabilities on the anatomical neighbors of the tracts using TRACULA, the anatomical segmentations were transformed to the subject's individual dMRI space. This within-subject, dMRI-to-T1 alignment was performed using a boundary-based, affine registration method (Greve & Fischl, 2009). To ensure that the relative position of the anatomical structures was the same for all subjects, and to map the median streamline from the training data to the subject during initialization, all subjects were mapped onto a template brain. We used a non-linear SyN registration in ANTs (Avants et al., 2008) to map each subject onto an FA template constructed from the training dataset (Maffei et al., 2021). It is important to highlight that this step is only used to initialize the reconstruction of the tract, which is then refined by fitting it to the anatomy of the individual subject.

2.6 Quality assessment of tract reconstructions

To assess the feasibility of implementing TRACULA in a clinical setting, we performed a qualitative assessment to evaluate the accuracy of the tract reconstructions in patients with acute severe TBI. We considered the application of TRACULA in this population to be feasible if more than 50% of the automated reconstructions were successful. For this qualitative assessment, a study investigator (C.M.) visually inspected all 40 tracts for all patients and healthy subjects. Reconstructions were considered successful if tracts traversed the relevant WM regions and reached the cortical regions that were used as inclusion ROIs in the protocols defined to manually label the training set. These protocols were carefully defined on the known anatomy of each tract and are detailed in (Maffei et al., 2021).

When a focal lesion was present, the reconstruction was considered successful if the tract reached the cortical termination regions delineated in (Maffei et al., 2021), even if the tract's course was altered by the presence of the lesion. As TRACULA uses a global probabilistic approach to estimate the tracts, post-processing steps to remove false positive reconstructions were not necessary. However, when a suitable solution for the initial reconstruction of the pathway is not found, the tract appears as a single curve (See Figure 1 for a simulated example). This can be due to errors in the subject-to-template

registration step, or to a misplacement of the FreeSurfer segmentation labels in the presence of lesions. We refer to these reconstructions as "failed reconstructions". Moreover, as a single threshold is used across all tracts and groups to threshold the posterior probability of the tract, some reconstructions can result in incomplete tracts or in unusually small tracts for which many voxels did not survive the threshold (See Figure 1 for an example). We refer to these reconstructions as "partial reconstructions". We reran TRACULA on the tracts that resulted in either failed or partial reconstructions by reinitializing the control points of the initial spline, by setting the *reinit* parameter to 1 (Yendiki et al., 2011). We retained the re-initialized reconstructions if they resulted in correct tracts as defined at the beginning of this section. The same criteria were applied to patients and controls. This was the only manual intervention performed in the TRACULA pipeline. Tracts that had partial or failed reconstructions after reinitialization were excluded from subsequent analyses. We performed individualized assessment of TAI using the remaining tracts, including both the tracts that had been successfully reconstructed without need of reinitialization and the tracts that needed to be reinitialized (*Section 2.7*). We then compared the results when performing individualized assessment of TAI using only the tracts that had been successfully reconstructed without need of reinitialization. We used a Wilcoxon signed-rank test in Python 3 (SciPy Stats 1.3.1) to assess differences in the number of injured tracts per patient.

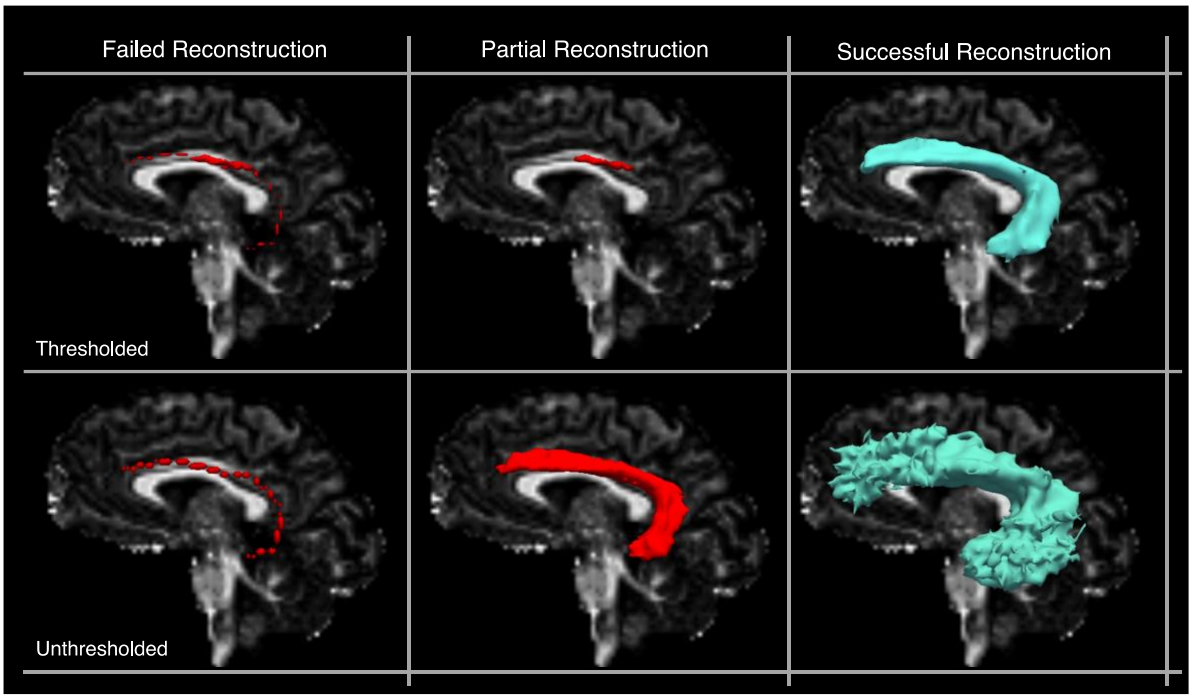

Figure 1. Simulated examples of failed, partial, and successful TRACULA reconstructions for a representative tract (the left arcuate fasciculus) in a control subject. Reconstructions are shown thresholded (20% of the maximum posterior probability) and unthresholded. Reconstructions are shown on sagittal fractional anisotropy scalar maps.

2.7 Multivariate analysis

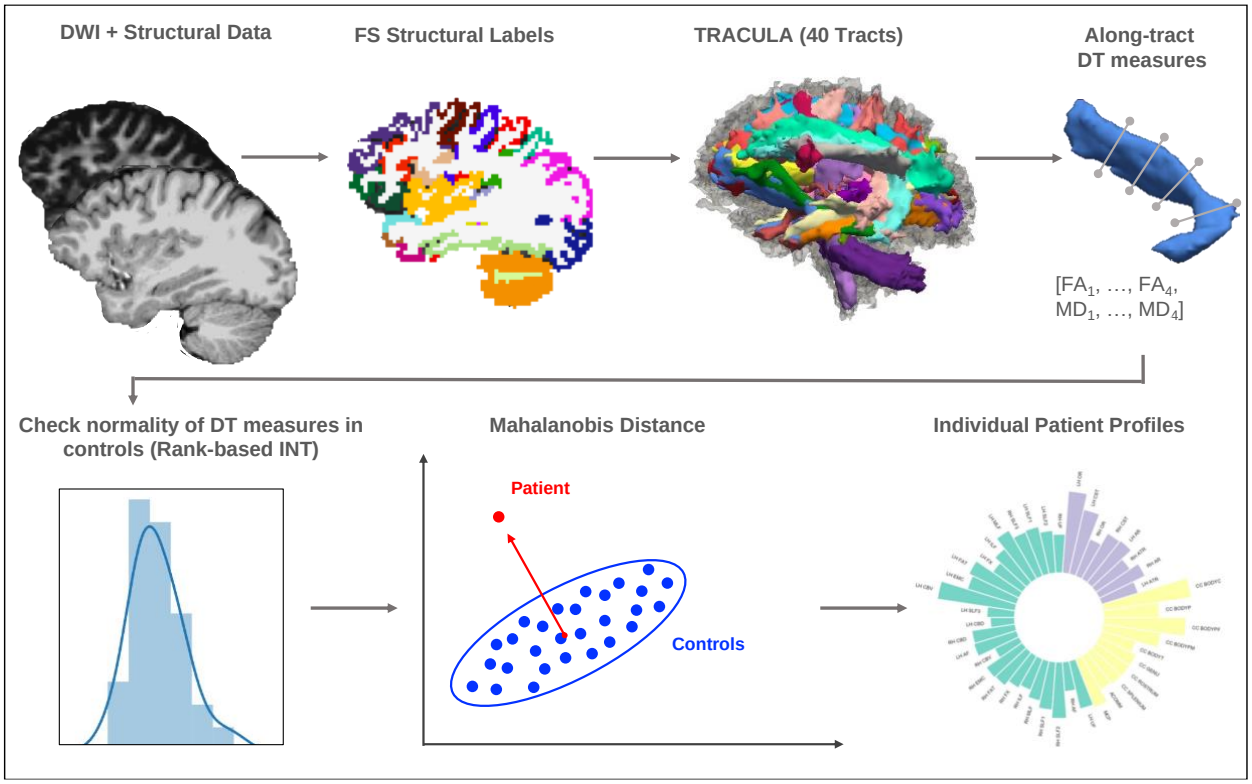

Figure 2. Schematic of the automated individualized pipeline. After pre-processing of diffusion weighted imaging (DWI) and structural T1-weighted imaging data, which included extraction of the FreeSurfer anatomical labels, the 40 white matter tracts were reconstructed in TRACULA. Along-tract measures derived from the diffusion tensor (DT) were then extracted for four equidistant segments for each of the 40 tracts. We checked the normality of the distribution for these measures in the controls and applied a Rank-based inverse normal transformation (INT) to the tracts for which assumptions of normality were not met ($p < 0.05$). We then computed the Mahalanobis distance for each tract using along-tract fractional anisotropy (FA) (FA_1, …, FA_4) and mean diffusivity (MD) (MD_1,…, MD_4) as features ($n = 8$). Individual profiles based on the distance of each tract from the controls were then built for each patient. FS: FreeSurfer.

Figure 2 shows the steps of the multivariate analysis. To measure the extent of WM injury in each patient, we employed a multivariate analysis that included along-tract measures of both FA and MD. For this analysis we used both the tracts that had been successfully reconstructed without need of reinitialization and the tracts that needed to be reinitialized (*Section 2.6*). For each subject, each tract was divided uniformly in $i = 1,..,4$ segments and the along-tract measures of FA and MD, extracted at each point along the tract in TRACULA (see *Section 2.4*), were averaged within each segment to obtain a total of eight features for each tract (FA_1, FA_2, FA_3, FA_4 , MD_1, MD_2, MD_3, MD_4). We decided to collapse the along-tract measurements in four segments to ensure that for each tract the number of features (p = 8) was lower than the number of controls (n = 33). To meet normality assumptions, we ran Shapiro-Wilk

tests in Python 3 using SciPy Stats 1.3.1 (Jones et al., 2001) and tested for the normality of FA_i and MD_i for each of the 40 tracts in the control group. For the tracts for which the distribution of FA_i or MD_i departed significantly from normality ($p < 0.05$) a rank-based inverse normal transformation (INT) was applied to the data for both controls and patients (Blom constant = $3 / 8$) (Blom, 1958) in Python 3, as in (Jolly et al., 2020).

We compared the 8-D feature vector of dMRI measures ($FA_{1,\ldots,4}$ and $MD_{1,\ldots,4}$) between each patient and the healthy population based on the Mahalanobis distance D_M, which is defined as (Mahalanobis, 1936):

$$D^2{}_M = (\mathbf{x} - \boldsymbol{\mu})^T \cdot \mathbf{C}^{-1} \cdot (\mathbf{x} - \boldsymbol{\mu}),$$

where $\mathbf{x}$ is the feature vector of a specific WM tract in a patient, $\boldsymbol{\mu}$ is the mean feature vector of the same tract over controls, and $\mathbf{C}$ is the covariance matrix of the feature vector in controls. The values of $D^2{}_M$ reflect how far a patient, represented as a point in the 8-D feature space, is from the normative distribution, which is estimated from the controls. The Mahalanobis distance has been used in neuroimaging studies, *e.g.*, to classify between patients with neurological diseases and controls (Lindemer et al., 2015) and to detect individual neurodevelopmental differences (Dean et al., 2017). We computed the D_M between patients and healthy controls for each of the 40 tracts to build individual profiles and identify the location and severity of WM injuries in patients. For multivariate Gaussian data, the distribution of $D^2{}_M$ values is known to be Chi-squared with degrees of freedom equal to the number of features (Gnanadesikan & Kettenring, 1972). We computed the p-values corresponding to the Chi-square statistic in Python 3 using SciPy Stats 1.3.1 (Jones et al., 2001) and considered a WM tract to be injured if its $p - value$ was < 0.001 ($p < 0.05$, Bonferroni adjusted for multiple comparisons across 40 tracts). We extracted the total number of injured tracts for each patient.

To measure the accuracy of the pipeline in the task of detecting TAI, we computed its performance in discriminating between patients and controls. We first computed the D_M between each control and the remaining control population in a leave-one-out fashion and then performed a receiver operating characteristic (ROC) analysis by thresholding both the alpha used to identify a tract as extreme ($thresholds = 0.001, 0.01, 0.05$) and the number of tracts used to assign a subject to the control or patient population ($range = 1{:}1{:}40$). For each combination of alpha and number of tracts we computed the false positives (FP; number of controls classified as patients), the false negative (FN; number of patients classified as controls), the true positives (TP; number of patients classified as patients), the

true negatives (TN; number of controls classified as controls), the true-positive rate (TPR; $TP/(TP + FN)$) and false-positive rate (FPR; $FP/(FP + TN)$). We obtained the ROC curve by plotting the TPR as a function of FPR and computed the area under the ROC curve (AUC) to quantify accuracy. To visualize the difference in ${D^2}_M$ values between the two populations we plotted the probability density functions of ${D^2}_M$ values for patients and controls for each of the 40 tracts. Finally, we compared the results when using only the tracts that had been successfully reconstructed without need of reinitialization (Section 2.6). We used a Wilcoxon signed-rank test in Python 3 (SciPy Stats 1.3.1) to assess differences in the number of injured tracts per patient.

2.8 Testing for the effect of head motion

To ensure differences in ${D^2}_M$ values between patients and controls were not influenced by head motion in the scanner, we first computed a total motion index (TMI) for each subject as described in (Yendiki et al., 2014). The TMI is a composite score of four motion measures: i) average volume-by-volume translation, ii) average volume-by-volume rotation, iii) percentage of slices with signal drop-out, and iv) signal drop-out severity. For each subject, the TMI was computed in the following manner:

$$TMI = \sum_{j=1}^{4} \frac{x_{mj} - M_j}{Q_j - q_j}$$

where $j = 1, \dots, 4$ are the four motion measures listed above, x_{mj} is the value of the *j*-th motion measure for the *m*-th subject and M_j, Q_j, and q_j are respectively the median, upper quartile, and lower quartile of the *j*-th measure over all the subjects (Yendiki et al., 2014). We then used a nonparametric Wilcoxon rank sum test in Python 3 (SciPy Stats 1.3.1) to assess if there were statistically significant differences in head motion between the patients and controls. We also performed a simple least squares regression in Python 3 using statsmodels 0.10.1 (Seabold & Perktold, 2010) to test if the interaction between group and TMI significantly predicted the ${D^2}_M$ for each of the 40 reconstructed tracts (Bonferroni adjusted alpha level $= 0.001$).

2.9 Testing the relationship with behavioral measures

Before undertaking our analyses, GCS total scores were calculated by summing the eye opening, motor, and verbal subscale scores. CRS-R total scores were derived by the summing the auditory function, visual function, motor function, oromotor/verbal function, communication, and arousal subscales scores.

Both the GCS and CRS-R were included in analyses as they have complementary strengths (*i.e.,* CRS-R has higher psychometric properties than the GCS; GCS is more widely-used and faster to administer than the CRS-R) (Bodien et al., 2021; Giacino et al., 2004; Teasdale et al., 2014).

Next, non-parametric Spearman correlations were conducted to test the relationship between our primary dependent variables (*i.e.,* CRS-R total, GCS total, days in coma, days to command-following), primary independent variables (*i.e.,* average $\mathrm{D_M}$ and number of injured tracts per patient), and potential confounding variables (*i.e.,* age, TMI, days to MRI) (Edlow et al., 2016). Results of these analyses informed variable selection for our subsequent analyses in two ways: i) only independent and confounding variables that were significantly correlated with the dependent variables were included in the linear regression model(s); and ii) if our dependent variables were significantly correlated with each other then, separate linear regression models were performed to avoid multicollinearity. Correlations were conducted in R using 'cor.test' (RStudio Team, 2020). Significance was set at $\alpha = 0.05$ to ensure identification of potential confounding variables.

Based on the results of the previously described correlations, linear regression analyses would be conducted using the average $\mathrm{D_M}$, the number of injured tracts, and any identified confounding variables to predict i) GCS Total scores, ii) CRS-R Total scores, ii) days in coma, and iv) days to command-following. Linear regressions were computed in R using the 'lm' function (RStudio Team, 2020). We also performed ordinal logistic regression using the average D_M of the injured tracts, the number of injured tracts, and any identified confounding variables to predict levels of consciousness (*i.e.*, levels from more to less severe impairment: coma, vegetative state, minimally conscious state, minimally conscious state plus, posttraumatic confusional state) (Giacino et al., 2004; Thibaut et al., 2020). This statistical approach was selected to manage the categorical nature of the level of consciousness variable. We tested the relationship between levels of consciousness and potential confounding variables (*i.e.,* days to MRI, age, and TMI), and included variables identified to be confounding in the final ordinal logistic regression model. Ordinal logistic regression analyses were computed in R (RStudio Team, 2020) using the 'plor' function from the "MASS" package (Venables & Ripley, 2002). T-values were compared against the standard normal distribution to obtain t-values for the estimates. To support interpretability, odds ratios and confidence intervals were obtained by exponentiating the estimates and confidence intervals

Then, as a secondary, exploratory analysis, ordinal logistic regression analyses were implemented using the average $\mathrm{D_M}$ and the number of affected tracts for the three tract subsets measuring motor (motor tracts (MT)), language comprehension (language comprehension tracts (LCT)), and language production (language production tracts (LPT)), and any identified confounding variables to predict the

corresponding GCS and CRS-R subscales as shown in Table 2. Tract subsets were chosen based on prior literature (Catani & Mesulam, 2008; Dick et al., 2019; Herbet et al., 2018; Maffei et al., 2019; Makris et al., 2005, 2009; Mars et al., 2016; Welniarz et al., 2017). Ordinal logistic regression was selected to manage the categorical nature of the GCS and CRS-R subscales (*i.e.,* see Table 1 for details). Non-parametric Spearman correlations were conducted in R to test the relationship between our continuous independent variables (*i.e*., the average D_M and number of injured tracts for the motor, language comprehension, and language production tract subsets) and potential confounding variables (*i.e.,* age, TMI, days to MRI). Ordinal logistic regression analyses were applied to test the relationship between the categorical dependent variables (*i.e*., GCS and CRS-R subscales) and potential confounding variables (i.e., days to MRI, age, and TMI). Independent variables or confounding variables that were significantly correlated with one another were included in separate models to manage multicollinearity. We focused these exploratory analyses on language and motor function (*e.g.,* as opposed to visual function) as their presence is considered a positive sign of recovery (Formisano et al., 2004; Tamashiro et al., 2012), reflecting reemergence of consciousness (Giacino et al., 2014). We included the CRS-R-A (*i.e*., assesses basic language comprehension) in addition to the CRS-R-O/V and GCS-V (*i.e*., assess basic language production) as these measures would be anticipated to be related to damage to different tracts, the ventral and dorsal streams, respectively (Hickok & Poeppel, 2004; Saur et al., 2008).

Domain	**Tract**	**Subscale**		**Reference for Tract Selection**
		Name	***Levels***	
Motor	Corticospinal tract (CST)	GCS-Motor	1 = No Response 2 = Abnormal extension 3 = Abnormal Flexion 4 = Flexion withdrawal from pain 5 = Moves to localized pain 6 = Obeys commands	Welniarz et al., 2017
		CRS-R Motor	0 = None 1 = Abnormal Posturing 2 = Flexion Withdrawal 3 = Localization to Noxious Stimulation 4 = Object Manipulation 5 = Automatic Motor Response 6 = Functional Object Use	
Language production	Frontal aslant tract (FAT)	GCS-Verbal	1 = No response 2 = Incomprehensible Sounds 3 = Incomprehensible Words 4 = Confused 5 = Oriented to time, place, and person	Dick et al., 2019
	Superior longitudinal fasciculus III (SLF3)	CRS-R Oromotor/ Verbal	0 = None 1 = Oral Reflexive Movement 2 = Vocalization/Oral Movement 3 = Intelligible Vocalization	Makris et al., 2015, 2009
	Arcuate fasciculus (AF)			Catani & Mesulam, 2008

Language comprehension	Arcuate fasciculus (AF)	CRS-R Auditory	0 = None 1 = Auditory Startle 2 = Localization to Sound 3 = Reproducible Movement to Command 4 = Consistent Movement to Command	Catani & Mesulam, 2008
	Acoustic radiation (AR)			Maffei et al., 2019
	Middle longitudinal fasciculus (MdLF)			Makris et al., 2015, 2009
	Extreme capsule (EMC)			Mars et al., 2016
	Inferior longitudinal fasciculus (ILF)			Herbet et al., 2018

Table 1. **Tract Selection for Correlation analysis.** The table shows the white matter tracts selected to perform correlations between the output of the multivariate analysis and the sub-scales of the CRS-R and GCS. For each domain assessed in the subscale, a group of tracts was selected based on prior literature. For each tract we selected its reconstruction in both hemispheres. The *levels* column denotes how the levels of the CRS-R and GCS subscales were ranked for use in the ordinal logistic regressions. CRS-R: Coma Recovery Scale-Revised; GCS: Glasgow Coma Scale.

3. Results

3.1 Patient Demographics and Clinical Characteristics

Patient demographics and clinical characteristics are provided in Table 2, including clinical indicators of coma duration (*i.e.,* days in coma, days until command-following, level of consciousness at MRI); and behavioral measures of consciousness (*i.e.,* GCS and CRS-R total scores), including subscales measuring language and motor function.

ID	Age Range	Sex	Days in Coma Range	Days Until Command - Following	Days to MRI Range	LOC at MRI	GCS-T at MRI	GCS-V at MRI	GCS-M at MRI	CRS-R-T at MRI	CRS-R-M at MRI	CRS-R-A at MRI	CRS-R-O/V at MRI
P1	26-30	M	1-5	1	16-20	PTCS	15	5	6	23	6	4	3
P2	21-25	M	1-5	6	1-5	MCS-	7	1	5	4	3	0	1
P3	16-20	F	1-5	10	1-5	Coma	5	1	3	1	1	0	0
P4	16-20	M	1-5	2	16-20	PTCS	14	4	6	23	6	4	3
P5	31-35	M	6-10	26	11-15	VS	7	3	1	3	0	0	2
P6	26-30	F	1-5	9	6-10	VS	9	1	4	6	2	0	1
P7	41-45	M	1-5	6	11-15	MCS+	13	3	6	18	5	3	3
P8	31-35	M	1-5	7	6-10	PTCS	11	1	6	20	5	4	2
P9	31-35	M	1-5	8	11-15	MCS+	10	1	6	9	2	3	1
P10	21-25	M	1-5	7	11-15	MCS-	10	1	5	10	5	1	1
P11	21-25	F	1-5	9	11-15	PTCS	14	4	6	22	6	4	3
P12	26-30	F	11-15	**	6-10	Coma	5	1	3	1	1	0	0
P13	16-20	M	1-5	2	1-5	MCS+	10	1	6	12	5	3	1
P14	51-55	M	6-10	**	6-10	VS	6	1	3	3	1	0	0
P15	26-30	M	1-5	8	6-10	MCS-	7	1	5	3	3	0	0
P16	31-35	M	1-5	3	1-5	MCS+	10	1	6	12	5	4	0

P17	26-30	F	1-5	**	11-15	VS	8	1	3	3	1	0	0
P18	26-30	M	1-5	0-2*	26-30	PTCS	11	1	6	18	6	3	2
Mean (SD)	28.56 (8.71)		3.00 (3.25)	7.00 (6.09)	10.39 (6.46)		9.56 (3.09)	1.78 (1.35)	4.78 (1.56)	10.61 (8.13)	3.50 (2.15)	1.83 (1.82)	1.28 (1.18)
Range	18-51		1-13	1-26	1 -28		5 - 15	1 - 5	1 - 6	1 - 23	0 - 6	0 - 4	0 - 3

Table 2. **Demographics and Clinical Characteristics of the TBI patients.** CRS-R: Coma Recovery Scale-Revised; CRS-R-T: Total score of the CRS-R; CRS-R-M: Motor CRS-R sub-scale; CRS-R-A: Auditory CRS-R sub-scale; CRS-R-O/V: Oromotor-verbal CRS-R sub-scale; GCS: Glasgow Coma Scale; GCS-T: Total score of the GSC; GCS-V: Verbal GCS sub-scale; GCS-M: Motor GCS sub-scale; LOC: level of consciousness; MCS+ = minimally conscious state plus, evidence of language function (*e.g.,* reproducible movement to command, object recognition, intelligible verbalization); MCS- = minimally conscious state minus, no evidence of language function, but shows at least one behavior consistent with a minimally conscious state (*e.g.,* automatic motor response). PTCS: posttraumatic confusional state; SD: standard deviation; VS: vegetative state. *Unable to determine exact date based on intensive care unit clinician assessments. We used 1 day for further analysis. **Patients who died in the ICU after withdrawal of life sustaining therapy and before command-following was observed.

3.2 Quality assessment and feasibility of tractography reconstructions

Figure 3 shows the reconstruction of the 40 tracts for six representative patients with different mechanisms of injury and one representative control. We observed that the TRACULA pipeline successfully reconstructed 93% of the tracts (mean: 37,5; SD: 2.18) across all patients without manual intervention. This is 43% more successful reconstructions than the threshold we set to consider the TRACULA automated stream feasible in this study (Section 2.6). Figure 4 reports the number of reconstructions for each patient that i) were correctly reconstructed without the need for reinitialization, ii) resulted in either partial or failed reconstruction and were reinitialized successfully, and iii) resulted in partial or failed reconstructions even after reinitialization (See *Section 2.6*). For a complete list of tracts that were reinitialized for each patient see supplementary table 1.

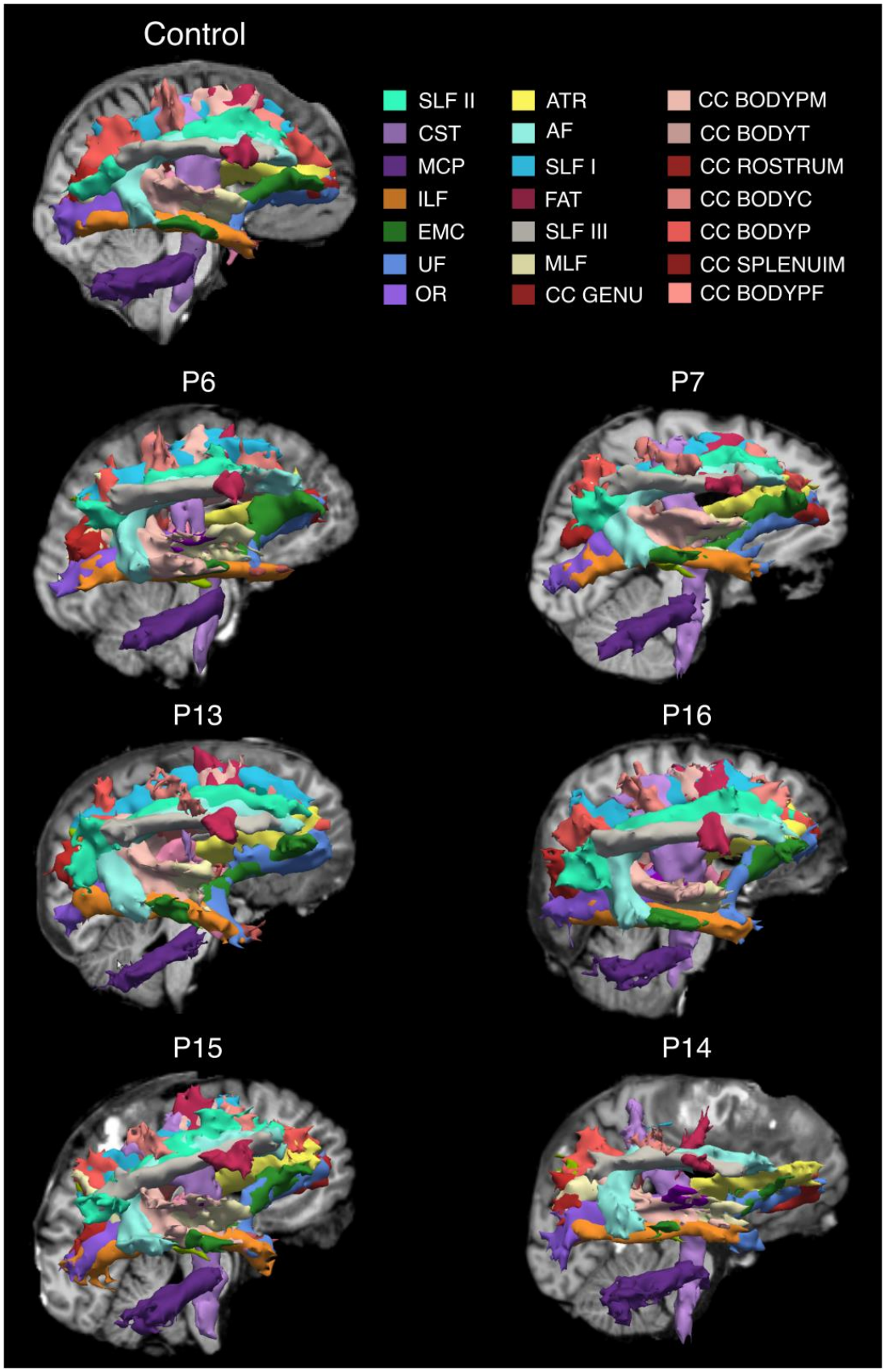
Control
SLF II
CST
MCP
ILF
EMC
UF
OR
ATR
AF
SLF I
FAT
SLF III
MLF
CC GENU
CC BODYPM
CC BODYT
CC ROSTRUM
CC BODYC
CC BODYP
CC SPLENUIM
CC BODYPF
P6
P7
P13
P16
P15
P14

Figure 3. TRACULA reconstructions. The 40 tracts automatically reconstructed in TRACULA are shown in 3D sagittal view (right hemisphere) for one control and six representative patients. For each subject, tracts are overlaid on the T1-weighted structural volume. Patients with few or no tracts visually affected by the lesion **(LEFT COLUMN**) and with more affected tracts (**RIGHT COLUMN**) are shown. Color labels are reported only for tracts visible in the figure. **Complete list of WM tracts abbreviations:** ACOMM: anterior commissure; AF: arcuate fasciculus; AR: acoustic radiation; ATR: anterior thalamic radiation; CC BODYC: central section of the body of the CC; CC BODYP: parietal section of the body of the CC; CC BODYPF: prefrontal section of the body of the CC; CC BODYPM: premotor section of the body of the CC; CC BODYT: temporal section of the body of the CC ; CC GENU: genu of the CC; CC ROSTRUM: rostrum of the CC; CC SPLENIUM: splenium of the CC; CBD: dorsal portion of cingulum bundle; CBV: ventral portion of the cingulum bundle; CC: corpus callosum; CST: cortico-spinal tract; EMC: extreme capsule; FAT: frontal Aslant tract; ILF: inferior longitudinal fasciculus; LH: left hemisphere; ILF: middle longitudinal fasciculus; OR: optic radiation; RH: right hemisphere; SLF I,II,III: first, second, and third branch of the superior longitudinal fasciculus; UF: uncinate fasciculus.

In one patient (P1) the FreeSurfer "recon-all" stream failed due to the presence of a massive lesion that injured a substantial portion of the cerebral cortex in the right hemisphere. This patient was thus excluded from the TRACULA pipeline. For five of the remaining 17 patients (P2, P3, P6, P12, P16), reconstructions were complete and accurate and did not require reinitialization, even in the presence of focal lesions (P2, P6, P12) (Supplementary Figure S1). For the remaining twelve patients, a small number of tracts (mean = 2.47, range = 1:7) resulted in failed or partial reconstructions and required reinitialization. For nine patients out of twelve, reinitialization led to successful reconstructions (as defined in Section 2.6). For one of these twelve patients (P4) two tracts were only partially reconstructed after reinitialization (MCP, LH CST), and for two patients (P10, P14) two and three tracts respectively (P10: ACOMM, RH UF; P14: ACOMM, RH SLFII, CC BODYPF) did not improve after reinitialization and resulted in failed reconstructions (Figure 3, P14). Upon visual inspection, failed or partial reconstructions were attributable to hemorrhagic contusions in the temporal and frontal lobes, respectively, which caused disruption of three tracts each (Supplementary Figure S2).

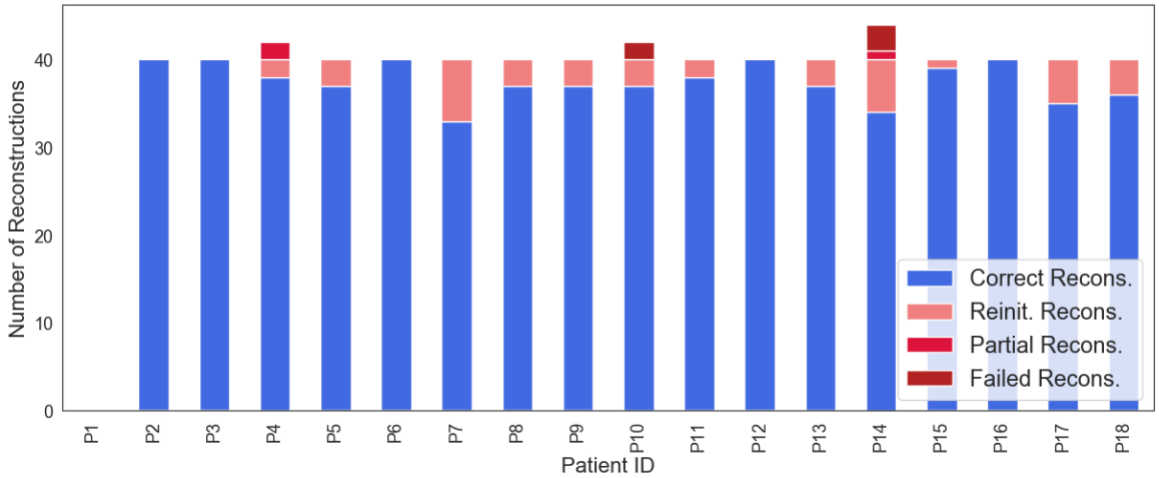

Figure 4. The number of correct reconstructions that did not need to be reinitialized, the number of reconstructions that were reinitialized, and number of tracts that resulted in partial or failed reconstructions even after reinitialization are shown for each of the patients TBI. The FreeSurfer automated segmentation pipeline failed in P1 due to the presence of a massive lesion. This precluded to run the TRACULA pipeline on this subject.

3.3 Multivariate individualized assessment of white matter damage

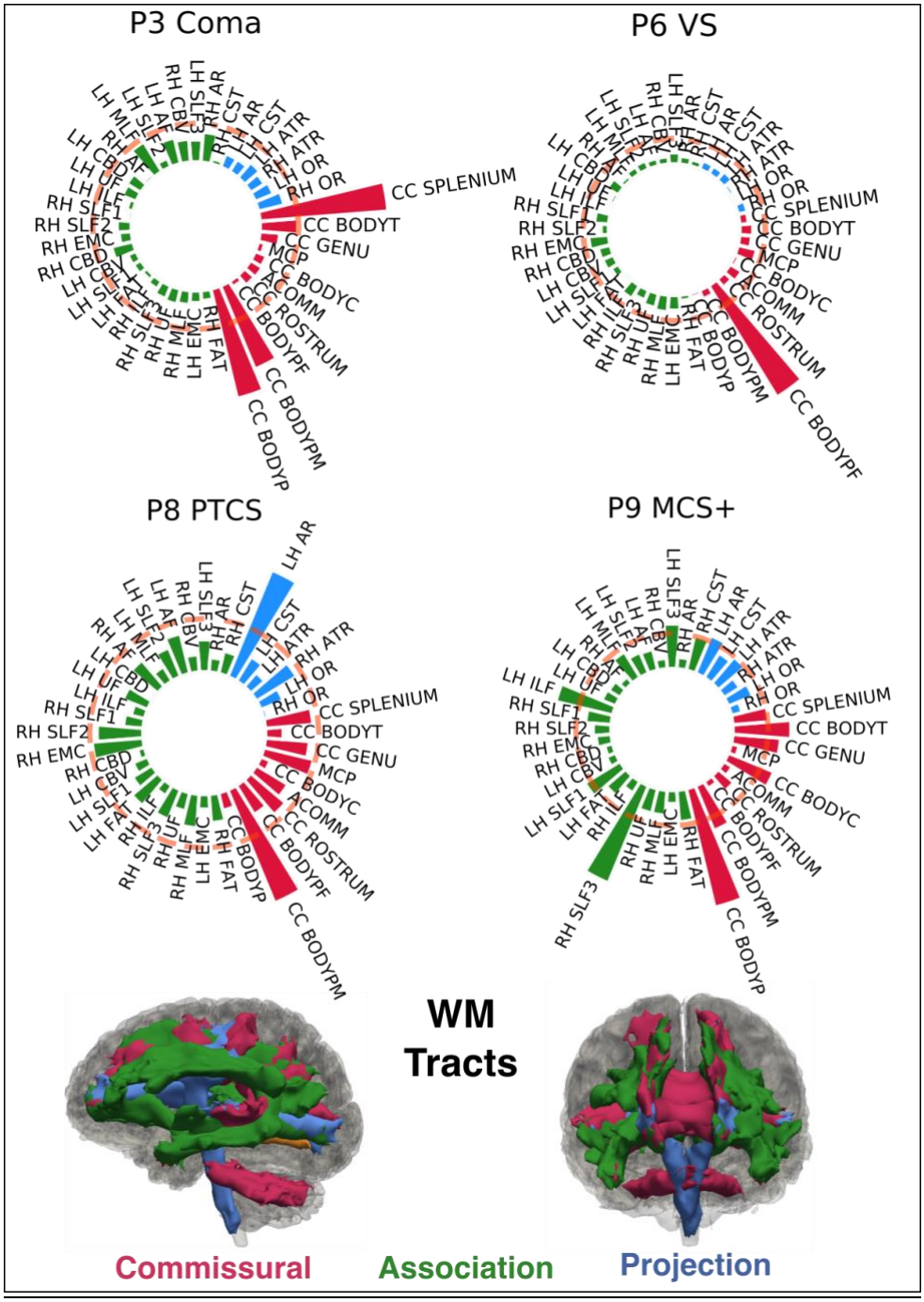

Figure 5. Individual Tract Profiles. The figure shows the individual profiles for four representative subjects, each with a different level of consciousness (Coma, VS, MCS+, PTCS). The Mahalanobis distance for the reconstructed tracts is plotted using polar bar plots. The pink dotted line corresponds to the critical value of the Chi-squared statistic at $\alpha = 0.001$. White matter tracts are grouped in commissural (red), associative (green), and projection pathways (blue). The 3D reconstructions of the 40 WM tracts are shown for a representative control subject, colored based on the group they belong to. MCS+: minimally conscious state plus; PTCS: posttraumatic confusional state; VS: vegetative state; WM: white matter. For a complete list of WM tract abbreviations see Figure 3.

The Shapiro-Wilk test showed that the distribution of FA_i and MD_i departed significantly from normality (p-value < 0.05) in 8.6 tracts per measure on average. Supplementary Table 2 reports the tracts along with the W- and p-value. Figure 5 shows the ${D^2}_M$-based individual profiles for four representative patients at different levels of consciousness (Coma, VS, MCS, PTCS). Individual profiles for all 17 patients are shown in Supplementary Figures S3 and S4. The individual profiles show the distance of each of the WM tracts from the multivariate distribution of that tract in the control population. To investigate whether a common pattern of TAI localization was visible across individuals, we grouped the reconstructed WM tracts as follows: i) commissural tracts: ACOMM, CC-BODYC, CC-BODYT, CC-BODYPF, CC-BODYP, CC-BODYPM, CC-GENU, CC-ROSTRUM, CC-SPLENIUM; ii) projection tracts: CST, ATR, AR, OR; iii) association tracts: all the other WM tracts. Commissural tracts were shown to have overall higher ${D^2}_M$ values (Supplementary Figure S5) that resulted in a higher number of significantly extreme values ($p < 0.001$) across patients compared to projection and association tracts (Supplementary Figure S6). The number and type of tracts that resulted injured varied across level of consciousness (LOC) (Figure 5, Figures S3, S4). For each patient we extracted the total number of injured tracts (we considered a WM tract to be injured if its $p - value$ was < 0.001 (Section 2.7)). On average 12.6 tracts resulted injured in patients with TBI (Figure 6).

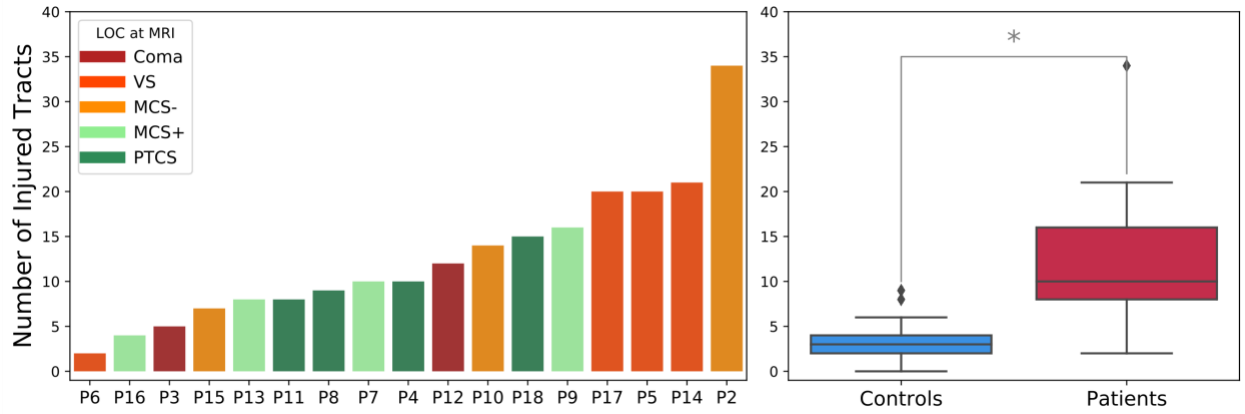

Figure 6. **Number of affected tracts per subject. Left panel)** The number of extreme WM tracts is shown for each subject with TBI. A tract was considered injured if its $p - value$, correspondent to a Chi-square statistic, was < 0.001 (p < 0.05 corrected for multiple comparisons across 40 tracts). Patients are ordered based on the number of injured tracts. The color of the bars reflects the level of consciousness (LOC) at the day of MRI. MCS+: minimally

conscious state plus; MCS-: minimally conscious state minus; PTCS: posttraumatic confusional state; VS: vegetative state. Patient IDs corresponding to those used in Table 1 ($P2, \ldots, P18$) are indicated for each patient. **Right panel)** Boxplots showing the number of extreme tracts for controls and patients. The asterisk indicates a significant difference (Wilcoxon sum rank test: $W = 4.79,\ p = 1.64e-6$).

To investigate the accuracy of the TRACULA pipeline at detecting TAI, we quantified its performance in classifying patients versus controls. We first computed the D_M for each control in a leave-one-out fashion and extracted the number of injured tracts for each control. On average 3.3 tracts resulted injured in controls (Figure 6, right panel). A Wilcoxon sum rank test showed significant difference between the number of injured tracts in patients and controls ($W = 4.79,\ p = 1.64e-6$). We then performed a ROC analysis and found an accuracy of 0.91 (Supplementary Figure 8). To visualize the difference in ${D^2}_M$ values between the two populations we plotted the probability density functions of ${D^2}_M$ for patients and controls. Figure 7 shows the probability density functions of ${D^2}_M$ values for the different sections of the corpus callosum for patients and controls. For subjects with severe TBI, ${D^2}_M$ values are shifted to the right suggesting a larger mean and greater degree of microstructural deviation from the normative reference group. Probability density functions of ${D^2}_M$ values for patients and controls for all 40 tracts are shown in supplementary figure S7.

Finally, there was no significant difference in number of injured tracts per patient when using only the tracts that had been successfully reconstructed without need of reinitialization (*i.e.*, discarding the tracts that needed to be reinitialized (*p=0.083*), suggesting the low number of tracts that needed to be reinitialized across subjects did not significantly impact the overall individualized assessment of TAI.

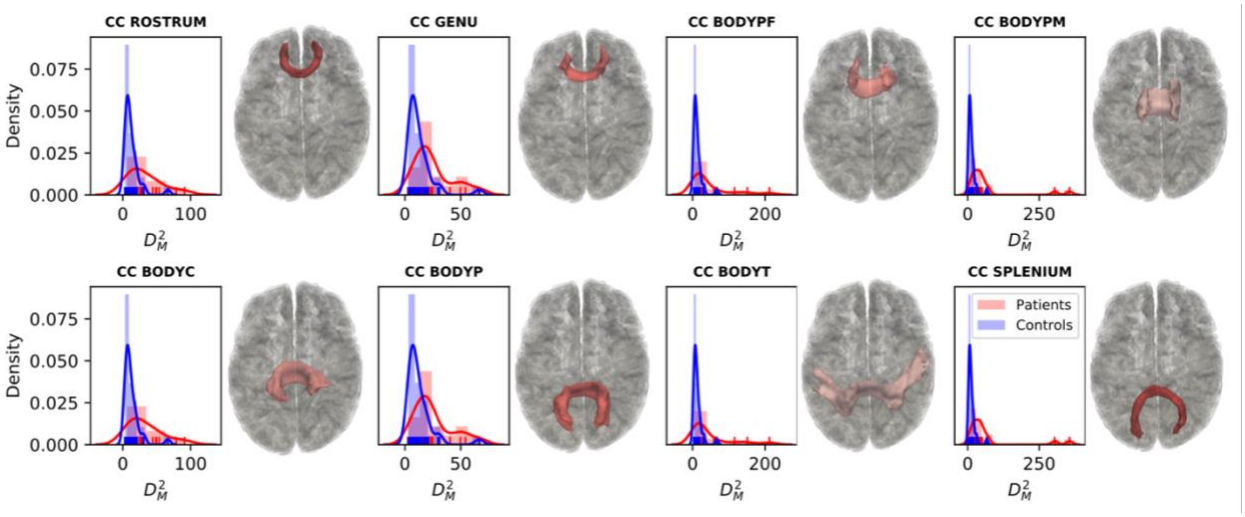

Figure 7. Density plots. Probability density functions of Mahalanobis distance (${D^2}_M$) values are shown for patients (shown in red) and controls (shown in blue) for each of the different sections of the corpus callosum (CC). Representative 3D reconstructions of the correspondent WM tracts are shown to the right of the density plots. CC

BODYC: central section of the body of the CC; CC BODYP: parietal section of the body of the CC; CC BODYPF: prefrontal section of the body of the CC; CC BODYPM: premotor section of the body of the CC; CC BODYT: temporal section of the body of the CC; CC GENU: genu of the CC; CC ROSTRUM: rostrum of the CC; CC SPLENIUM: splenium of the CC

3.4 Effects of head motion

The TMI and the four motion measures used to compute it are reported for each subject in Supplementary Table S4. A nonparametric Wilcoxon sum rank test revealed a non-statistically significant difference in TMI between patients and controls ($W = 1.91$, $p = 0.055$) and the group-by-TMI interaction did not significantly predict ${D^2}_M$ values for any of the 40 tracts nor the number of injured tracts per subject (all $p - values > 0.05$). Taken together, these results emphasize that the greater number of injured tracts detected for patients versus controls was not confounded by differences in head motion between the groups.

3.5 Relationship with behavioral measures

Non-parametric Spearman correlations identified several significant relationships as shown in Supplementary Table S4. See Supplementary Figure S10 for full results of the Spearman correlations testing associations between continuous variables.

Based on these findings, separate linear regression analyses were conducted for each of the dependent variables (*i.e.,* GCS-Total, CRS-R Total, Days in Coma, Days to command-following) with both the average D_M and number of affected tracts included as predictor variables. Days to MRI was included as a continuous covariate in the regression predicting GCS-Total, as they were revealed to be significantly correlated (*rho* = 0.54, $p < 0.01$). Average D_M and number of affected tracts did not significantly predict GCS-Total, CRS-R Total, Days to command-following, or Days in coma (all uncorrected *p-values* > 0.05). Days to MRI did not significantly predict GCS-Total in the regression model (*beta* = 0.21, *SE* = 0.10, t = 2.048, $p = 0.06$). No confounding variables were identified for the ordinal regression using the average D_M and number of affected tracts to predict levels of consciousness, although Days to MRI approached significance (*beta* = 0.37, *SE* = 0.20, t = 1.87, $p = 0.06$). Average D_M and number of affected tracts did not significantly predict level of consciousness (*beta* = -0.01, *SE* = 0.02, $t = -0.73$, $p = 0.46$ and beta = -0.03, *SE* = 0.09, $t = -0.35$, $p = 0.72$, respectively).

In terms of our secondary, exploratory analyses with language and motor function, non-parametric Spearman correlations identified several relationships between our dependent (i.e., GCS Motor, GCS Verbal, CRS-R-Auditory, CRS-R-Oromotor/Verbal, CRS-R Motor) and independent variables

(*e.g.,* average D_M of tract subsets, n of affected tracts within each subset) as shown in Supplementary Table S4. These results influenced our analyses in two ways: i) separate ordinal regression analyses were undertaken for each of the dependent variables (*i.e.,* GCS Verbal, GCS Motor, CRS-R Auditory, CRS-R-Oromotor/Verbal, CRS-R-R Motor); and ii) the average D_M of tract subsets and the number of affected tracts within a subset were not included in the same model. Before conducting ordinal logistic regressions testing the relationship between the dMRI-derived measures and language and motor subscales of the GCS and CRS-R, the relationship between age, TMI, and days to MRI and the sub-scales were individually investigated. TMI (*beta* = 0.11, *SE* = 0.05, t = 1.98, *p* = 0.047) showed a significant relationship with the GCS Verbal subscale and thus, was included as a confounding variable. TMI (*beta* = 0.14, *SE* = 0.06, *t* = 2.39, *p* = 0.167) and Days to MRI (*beta* = 0.16, *SE* = 0.08, *t* = 1.98, *p* = 0.047) demonstrated significant relationships with the CRS-R Oromotor/Verbal subscale and thus, were included as confounding variables.

Neither the average D_M, nor the number injured tracts for the motor, language comprehension, and language production subsets significantly predicted their corresponding GCS and CRS-R subscales (all uncorrected $p - values > 0.05$). Interestingly, however, for every one unit increase in TMI, the likelihood of a patient demonstrating a higher level of verbal expression function (as measured by the GCS Verbal subscale) increased by 18% in the ordinal logistic regression with the average D_M of language production tracts as a predictor variable (*beta* = 0.16 , *SE* = 0.09, *t* = 1.83 , *p* = 0.06, *Profiled 95% Confidence Interval* = 0.02 to 0.38, *Proportional Odds Ratio* = 1.18) and by 19% in the ordinal logistic regression with the number of affected language production tracts as a predictor variable (*beta* = 0.17, *SE* = 0.09 , *t* =1.91 , *p* = 0.055, *Profiled 95% Confidence Interval* = 0.03 to 0.38, *Proportional Odds Ratio* = 1.19).

Furthermore, for every one unit increase in days to MRI, the likelihood of a patient demonstrating a higher level of verbal expression function (as measured by the CRS-R Oromotor/Verbal subscale) increased by 17% in the ordinal logistic regression with the average D_M of language production tracts as a predictor variable (*beta* = 0.16 , *SE* = 0.08, *t* = 1.98, *p* = 0.04, *Profiled 95% Confidence Interval* = 0.02 to 0.34, *Proportional Odds Ratio* = 1.17) and by 18% in the ordinal logistic regression with the number of affected language production tracts as a predictor variable (*beta* = 0.16, *SE* = 0.08 , *t* = 1.99, *p* = 0.045, *Profiled 95% Confidence Interval* = 0.02 to 0.35, *Proportional Odds Ratio* = 1.18).

Finally, for every one unit increase in TMI, the likelihood of a patient demonstrating a higher level of verbal expression function (as measured by the CRS-R Oromotor/Verbal subscale) increased by 14 % in the ordinal logistic regression with the average D_M of language production tracts as a predictor variable (*beta* = 0.13, *SE* = 0.06 , *t* = 2.39, *p* = 0.017, *Profiled 95% Confidence Interval* = 0.03 to 0.26, *Proportional Odds Ratio* = 1.14) and by 16 % in the ordinal logistic regression with the number of affected language

production tracts as a predictor variable (*beta* = 0.15 , *SE* = 0.06, *t* = 2.48, *p* = 0.013, *Profiled 95% Confidence Interval* = 0.04 to 0.28, *Proportional Odds Ratio* = 1.16).

All statistical analyses were conducted again using data from the completed tracts only. There were no differences in the results using the D_M or number of affected tracts that had been successfully reconstructed without need of reinitialization. This finding further supports that the few tracts that needed to be reinitialized across subjects did not significantly impact the overall individualized assessment of TAI.

4. Discussion

In this prospective observational study, we show that TRACULA provides automated reconstruction of white matter tracts from diffusion MRI data in critically ill patients with acute severe TBI, even in the presence of large focal lesions. These findings demonstrate the feasibility of applying the TRACULA automated tractography pipeline (Maffei et al., 2021; Yendiki et al., 2011) during the acute stage of severe TBI, addressing a major barrier to the assessment of TAI in this population. If replicated in larger studies, these proof-of-principle results suggest that TRACULA can be used by clinicians for acute detection of TAI in critically ill patients – information that can inform prognosis, guide therapeutic decision-making, and provide additional information to the standardized clinical assessment. We discuss the clinical implications and technical limitations of the current results and provide further directions for optimization of the proposed pipeline.

With respect to TAI detection, TRACULA provided individualized profiles of WM disruption that differentiated patients from healthy controls and from one another (Fig. 5, Fig. 6), providing an acute assessment of TAI burden for each individual patient. This finding builds on recent studies showing that DT measures can be used to identify the neuroanatomic distribution and severity of TAI at the individual level (Jolly et al., 2020), an important goal given the well-established heterogeneity of this condition (Douglas et al., 2019). Multivariate approaches that combine different neuroimaging measures have shown higher power in distinguishing different clinical populations from controls (Dean et al., 2017). Indeed, the multivariate analysis of along-tract DT metrics revealed marked variability in the number and location of injured WM tracts between patients (Fig. 5, Fig. 6). While no clear pattern of WM damage was visible within LOC, results of the multivariate individual analysis revealed commissural connections to be overall more significantly and more frequently affected compared to projection and association fibers (Fig. 5, Fig. S5, Fig. S6). Specifically, the frontal, parietal, and temporal sections of the CC were the most affected, consistent with previous histological and neuroimaging studies in humans (Moen et al., 2014;

Nolan et al., 2020; Ubukata et al., 2016) and with studies in animal models of TAI (Baker et al., 2004; Gennarelli et al., 1982).

Interestingly, individual multivariate measures of TAI did not show a significant relationship with behavioral measures of consciousness (i.e., GCS-Total, CRS-R Total, days to command-following, days in coma) or basic language and motor function (*i.e.,* GCS Verbal, GCS Motor, CRS-R Auditory, CRS-R Oromotor/Verbal, CRS-R Motor). There are several potential explanations for this observation. First, widespread tract disruption in the brainstem may have confounded any associations between i) the global burden of TAI and behavioral measures of consciousness; and ii) TAI within domain-specific tracts and behavioral measures of language or motor function. TRACULA does not yet provide automated reconstructions of brainstem tracts – a key direction for future work in this field, especially as the presence of TAI in the brainstem pathways can cause altered consciousness, and impaired language and motor function (Edlow et al., 2013; Snider et al., 2019). Second, it is important to acknowledge that the GCS and CRS-R subscales are not "pure" measures of motor and language function (*e.g.*, highest score on the GCS motor requires intact auditory comprehension) (Schnakers et al., 2015), and these scores may not reflect "covert consciousness" in the presence of cognitive-motor dissociation (Schiff, 2015). As prior correlations between acute tractography data and long-term cognitive function indicate a possible prognostic role for tract-specific TAI assessments (Wang et al., 2008), it will be important in future work to assess if these multivariate tract-specific measurements in the ICU are relevant to overall and/or domain-specific recovery in the subacute-to-chronic stages of care, when language and motor assessments may be less affected by global injury and/or brainstem injury.

TRACULA allowed to automatically reconstruct on average 93% of the WM tracts with minimal manual intervention, providing potential for dissemination to a broad range of hospital settings without the need for local tractography expertise. However, the potential clinical translation of this pipeline for TAI detection in the ICU will depend upon additional methodological and logistical factors that require further study. While on average TRACULA could successfully reconstruct 37.5 of 40 tracts without the need of manual intervention, an average of 2.4 tracts resulted in failed or partial reconstructions and needed to be reinitialized in 12 patients (Supplementary Table 1). This step relied upon a visual quality assessment from a tractography expert that would limit its applicability to clinical settings. Furthermore, some of the reinitialized tracts could not be successfully reconstructed due to the presence of focal lesions in two patients. While it has been previously shown that TRACULA is robust to errors in the anatomical segmentation (Zöllei et al., 2019), because prior probabilities are based on the relative positions of the bundles with respect to the labels, large lesions can lead to missing labels that can affect tract

reconstructions. Further optimization of the TRACULA pipeline will include i) automated quality assessment of tract reconstructions and ii) correct handling of missing FreeSurfer anatomical labels, including in cases where the FreeSurfer "recon-all" pipeline fails due to the presence of lesions (*e.g.,* Patient 1 in this study). Additionally, the dMRI data in this study were acquired on a clinical scanner in a research hospital. More specifically, the dMRI sequence used high spatial resolution (2 mm isotropic) and high angular resolution ($b = 2{,}000\ s/mm^2$, 60 diffusion-encoding directions) leading to a longer acquisition time (~ 9 minutes), relative to more standard clinical diffusion sequences (Chilla et al., 2015). Thus, while obtained on a clinical scanner, the dMRI data collected in this study were likely of "higher-quality" than dMRI data acquired on clinical scanners in hospitals not associated with a research facility, and it will be important to determine whether automated reconstruction of WM tracts with TRACULA is robust in the setting of lower-resolution, lower-quality dMRI data in the future. It is also important to consider that the individualized assessment can only be performed comparing patients to a control cohort scanned with the same dMRI sequence. For this pipeline to be applied in a clinical setting that does not have data from healthy controls, future optimizations of this pipeline should explore its robustness to dMRI data acquired on a different scanner and with a different dMRI sequence. Finally, it is important to consider that some patients with acute severe TBI are not stable enough to travel to an MRI scanner, and thus this pipeline may not be feasible to use until these patients reach the subacute stage of care.

The results of this study should also be interpreted in the context of further methodological limitations, including the small sample size. Although there was no statistically significant difference in age between the patient and control cohorts, the controls were not enrolled with an even distribution of age, sex and educational attainment. Identification of optimal characteristics and sample sizes for control cohorts is an area of active inquiry (Jolly et al., 2020), and future studies may leverage large normative databases such as the Human Connectome Project in this effort (Bookheimer et al., 2019; Harms et al., 2018).

In summary, this prospective observational study demonstrates the feasibility and utility of an automated tractography pipeline for detecting TAI in the acute injury phase of severe TBI. Quantitative analysis of patient-specific TAI burden in the ICU could increase the accuracy of outcome prediction in this population and guide decisions regarding the provision of life-sustaining treatment and rehabilitation.

Funding:
This study was supported by the NIH National Institute of Neurological Disorders and Stroke (R01NS119911, R21NS109627, RF1NS115268, U01NS1365885, U01NS086090), NIH Director's Office (DP2 HD101400), NIH National Institute of Biomedical Imaging and Bioengineering (R01EB021265, U01EB026996), NIH National Institute of Mental Health (U01MH108168, R56MH121426), James S. McDonnell Foundation, American Academy of Neurology, Tiny Blue Dot Foundation, and National Institute on Disability, Independent Living and Rehabilitation Research (NIDILRR), Administration for Community Living (90DPCP0008-01-00, 90DP0039).

Acknowledgments:
We thank the nursing staffs of the Massachusetts General Hospital Neurosciences ICU, Multidisciplinary ICU, and Surgical ICU. We also thank the Massachusetts General Hospital MRI technologists for assistance with data acquisition. We are grateful to the patients and families in this study for their participation and support.

Disclosures:
The authors report no relevant disclosures.

References:

Adams, J. H., Doyle, D., Ford, I., Gennarelli, T. A., Graham, D. I., & Mclellan, D. R. (1989). Diffuse axonal injury in head injury: Definition, diagnosis and grading. *Histopathology*, *15*(1), 49–59. https://doi.org/10.1111/j.1365-2559.1989.tb03040.x

Andersson, J. L. R., & Sotiropoulos, S. N. (2015). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2015.10.019

Arfanakis, K., Haughton, V. M., Carew, J. D., Rogers, B. P., Dempsey, R. J., & Meyerand, M. E. (2002). Diffusion tensor MR imaging in diffuse axonal injury. *American Journal of Neuroradiology*, *23*(5), 794–802.

Asken, B. M., DeKosky, S. T., Clugston, J. R., Jaffee, M. S., & Bauer, R. M. (2017). Diffusion tensor imaging (DTI) findings in adult civilian, military, and sport-related mild traumatic brain injury (mTBI): a systematic critical review. *Brain Imaging and Behavior 2017 12:2*, *12*(2), 585–612. https://doi.org/10.1007/S11682-017-9708-9

Avants, B. B., Epstein, C. L., Grossman, M., & Gee, J. C. (2008). Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Baker, A. J., Phan, N., Moulton, R. J., Fehlings, M. G., Yucel, Y., Zhao, M., Liu, E., & Tian, G. F. (2004). Attenuation of the Electrophysiological Function of the Corpus Callosum after Fluid Percussion Injury in the Rat. *Https://Home.Liebertpub.Com/Neu*, *19*(5), 587–599. https://doi.org/10.1089/089771502753754064

Basser, P. J., Mattiello, J., & LeBihan, D. (1994). MR diffusion tensor spectroscopy and imaging. *Biophysical Journal*, *66*(1), 259–267. https://doi.org/10.1016/S0006-3495(94)80775-1

Behrens, T. E. J., Johansen-Berg, H., Woolrich, M. W., Smith, S. M., Wheeler-Kingshott, C. a M., Boulby, P. a, Barker, G. J., Sillery, E. L., Sheehan, K., Ciccarelli, O., Thompson, a J., Brady, J. M., & Matthews, P. M. (2003). Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, *6*(7), 750–757. https://doi.org/10.1227/01.NEU.0000309595.77090.89

Benson, R. R., Meda, S. A., Vasudevan, S., Kou, Z., Govindarajan, K. A., Hanks, R. A., Millis, S. R., Makki, M., Latif, Z., Coplin, W., Meythaler, J., & Haacke, E. M. (2007). Global White Matter Analysis of Diffusion Tensor Images Is Predictive of Injury Severity in Traumatic Brain Injury. *Https://Home.Liebertpub.Com/Neu*, *24*(3), 446–459. https://doi.org/10.1089/NEU.2006.0153

Betz, J., Zhuo, J., Roy, A., Shanmuganathan, K., & Gullapalli, R. P. (2012). Prognostic value of diffusion tensor imaging parameters in severe traumatic brain injury. In *Journal of Neurotrauma* (Vol. 29, Issue 7, pp. 1292–1305). J Neurotrauma. https://doi.org/10.1089/neu.2011.2215

Blom, G. (1958). *Statistical estimates and transformed beta-variables*. Wiley.

Blumbergs, P. C., Scott, G., Vis, J. M., Wainwright, H., Simpson, D. A., & Mclean, A. J. (1995). Topography of axonal injury as defined by amyloid precursor protein and the sector scoring method in mild and severe closed head injury. *Journal of Neurotrauma*, *12*(4), 565–572. https://doi.org/10.1089/NEU.1995.12.565

Bodien, Y. G., Barra, A., Temkin, N. R., Barber, J., Foreman, B., Vassar, M., Robertson, C., Taylor, S. R., Markowitz, A. J., Manley, G. T., Giacino, J. T., & Edlow, B. L. (2021). Diagnosing Level of Consciousness: The Limits of the Glasgow Coma Scale Total Score. *Journal of Neurotrauma*, *38*(23), 3295–3305. https://doi.org/10.1089/NEU.2021.0199

Bookheimer, S. Y., Salat, D. H., Terpstra, M., Ances, B. M., Barch, D. M., Buckner, R. L., Burgess, G. C., Curtiss, S. W., Diaz-Santos, M., Elam, J. S., Fischl, B., Greve, D. N., Hagy, H. A., Harms, M. P., Hatch,

O. M., Hedden, T., Hodge, C., Japardi, K. C., Kuhn, T. P., … Yacoub, E. (2019). The Lifespan Human Connectome Project in Aging: An overview. *NeuroImage*, *185*, 335–348. https://doi.org/10.1016/J.NEUROIMAGE.2018.10.009

Brody, D. L., Benetatos, J., Bennett, R. E., Klemenhagen, K. C., & Mac Donald, C. L. (2015). The pathophysiology of repetitive concussive traumatic brain injury in experimental models; new developments and open questions. *Molecular and Cellular Neuroscience*, *66*(PB), 91–98. https://doi.org/10.1016/J.MCN.2015.02.005

Catani, M., & Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. *Cortex*, *44*(8), 953–961. https://doi.org/10.1016/j.cortex.2008.04.002

Chilla, G. S., Tan, C. H., Xu, C., & Poh, C. L. (2015). Diffusion weighted magnetic resonance imaging and its recent trend—a survey. *Quantitative Imaging in Medicine and Surgery*, *5*(3), 407. https://doi.org/10.3978/J.ISSN.2223-4292.2015.03.01

Colby, J. B., Soderberg, L., Lebel, C., Dinov, I. D., Thompson, P. M., & Sowell, E. R. (2012). Along-tract statistics allow for enhanced tractography analysis. *NeuroImage*, *59*(4), 3227–3242. https://doi.org/10.1016/j.neuroimage.2011.11.004

Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., & Friston, K. (2007). Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *NeuroImage*, *37*(3), 866–875. https://doi.org/10.1016/J.NEUROIMAGE.2007.04.065

D'Souza, M. M., Trivedi, R., Singh, K., Grover, H., Choudhury, A., Kaur, P., Kumar, P., & Tripathi, R. P. (2015). Traumatic brain injury and the post-concussion syndrome: A diffusion tensor tractography study. *The Indian Journal of Radiology & Imaging*, *25*(4), 404. https://doi.org/10.4103/0971-3026.169445

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Dean, D. C., Lange, N., Travers, B. G., Prigge, M. B., Matsunami, N., Kellett, K. A., Freeman, A., Kane, K. L., Adluru, N., Tromp, D. P. M., Destiche, D. J., Samsin, D., Zielinski, B. A., Fletcher, P. T., Anderson, J. S., Froehlich, A. L., Leppert, M. F., Bigler, E. D., Lainhart, J. E., & Alexander, A. L. (2017). Multivariate characterization of white matter heterogeneity in autism spectrum disorder. *NeuroImage : Clinical*, *14*, 54. https://doi.org/10.1016/J.NICL.2017.01.002

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. https://doi.org/10.1016/j.neuroimage.2006.01.021

Dick, A. S., Garic, D., Graziano, P., & Tremblay, P. (2019). The frontal aslant tract (FAT) and its role in speech, language and executive function. In *Cortex* (Vol. 111, pp. 148–163). Masson SpA. https://doi.org/10.1016/j.cortex.2018.10.015

Douglas, D. B., Ro, T., Toffoli, T., Krawchuk, B., Muldermans, J., Gullo, J., Dulberger, A., Anderson, A. E., Douglas, P. K., & Wintermark, M. (2019). Neuroimaging of Traumatic Brain Injury. *Medical Sciences*, *7*(1), 2. https://doi.org/10.3390/MEDSCI7010002

Edlow, B. L., Chatelle, C., Spencer, C. A., Chu, C. J., Bodien, Y. G., O'Connor, K. L., Hirschberg, R. E., Hochberg, L. R., Giacino, J. T., Rosenthal, E. S., & Wu, O. (2017). Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain*, *140*(9), 2399–2414. https://doi.org/10.1093/brain/awx176

Edlow, B. L., Claassen, J., Schiff, N. D., & Greer, D. M. (2021). Recovery from disorders of consciousness: mechanisms, prognosis and emerging therapies. *Nature Reviews. Neurology*, *17*(3), 135. https://doi.org/10.1038/S41582-020-00428-X

Edlow, B. L., Copen, W. A., Izzy, S., Bakhadirov, K., van der Kouwe, A., Glenn, M. B., Greenberg, S. M.,

Greer, D. M., & Wu, O. (2016). Diffusion tensor imaging in acute-to-subacute traumatic brain injury: a longitudinal analysis. *BMC Neurology*, *16*(1). https://doi.org/10.1186/S12883-015-0525-8

Edlow, B. L., Haynes, R. L., Takahashi, E., Klein, J. P., Cummings, P., Benner, T., Greer, D. M., Greenberg, S. M., Wu, O., Kinney, H. C., & Folkerth, R. D. (2013). *Disconnection of the Ascending Arousal System in Traumatic Coma*. www.trackvis.org

Fan, Q., Witzel, T., Nummenmaa, A., Van Dijk, K. R. A., Van Horn, J. D., Drews, M. K., Somerville, L. H., Sheridan, M. A., Santillana, R. M., Snyder, J., Hedden, T., Shaw, E. E., Hollinshead, M. O., Renvall, V., Zanzonico, R., Keil, B., Cauley, S., Polimeni, J. R., Tisdall, D., … Rosen, B. R. (2015). MGH–USC Human Connectome Project datasets with ultra-high b-value diffusion MRI. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2015.08.075

Fischl, B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/J.NEUROIMAGE.2012.01.021

Formisano, R., Carlesimo, G. A., Sabbadini, M., Loasses, A., Penta, F., Vinicola, V., Caltagirone, C., & Wilson, L. (2004). Clinical predictors and neuropsychological outcome in severe traumatic brain injury patients. *Acta Neurochirurgica*, *146*(5), 457–462. https://doi.org/10.1007/S00701-004-0225-4

Galanaud, D., Perlbarg, V., Gupta, R., Stevens, R. D., Sanchez, P., Tollard, E., De Champfleur, N. M., Dinkel, J., Faivre, S., Soto-Ares, G., Veber, B., Cottenceau, V., Masson, F., Tourdias, T., André, E., Audibert, G., Schmitt, E., Ibarrola, D., Dailler, F., … Puybasset, L. (2012). Assessment of White Matter Injury and Outcome in Severe Brain TraumaA Prospective Multicenter Cohort. *Anesthesiology*, *117*(6), 1300–1310. https://doi.org/10.1097/ALN.0B013E3182755558

Gennarelli, T. A., Thibault, L. E., Adams, J. H., Graham, D. I., Thompson, C. J., & Marcincin, R. P. (1982). Diffuse axonal injury and traumatic coma in the primate. *Annals of Neurology*, *12*(6), 564–574. https://doi.org/10.1002/ana.410120611

Giacino, J. T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: the state of the science. *Nature Reviews Neurology 2014 10:2*, *10*(2), 99–114. https://doi.org/10.1038/nrneurol.2013.279

Giacino, J. T., Kalmar, K., & Whyte, J. (2004). The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. *Archives of Physical Medicine and Rehabilitation*, *85*(12), 2020–2029. https://doi.org/10.1016/J.APMR.2004.02.033

Gnanadesikan, R., & Kettenring, J. R. (1972). Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data. *Biometrics*, *28*(1), 81. https://doi.org/10.2307/2528963

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Harms, M. P., Somerville, L. H., Ances, B. M., Andersson, J., Barch, D. M., Bastiani, M., Bookheimer, S. Y., Brown, T. B., Buckner, R. L., Burgess, G. C., Coalson, T. S., Chappell, M. A., Dapretto, M., Douaud, G., Fischl, B., Glasser, M. F., Greve, D. N., Hodge, C., Jamison, K. W., … Yacoub, E. (2018). Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. *NeuroImage*, *183*, 972–984. https://doi.org/10.1016/J.NEUROIMAGE.2018.09.060

Hashim, E., Caverzasi, E., Papinutto, N., Lewis, C. E., Jing, R., Charles, O., Zhang, S., Lin, A., Graham, S. J., Schweizer, T. A., Bharatha, A., & Cusimano, M. D. (2017). Investigating microstructural abnormalities and neurocognition in sub-acute and chronic traumatic brain injury patients with normal-appearing white matter: A preliminary diffusion tensor imaging study. *Frontiers in Neurology*, *8*(MAR), 97. https://doi.org/10.3389/fneur.2017.00097

Herbet, G., Zemmoura, I., & Duffau, H. (2018). Functional Anatomy of the Inferior Longitudinal Fasciculus: From Historical Reports to Current Hypotheses. *Frontiers in Neuroanatomy*, *12*. https://doi.org/10.3389/FNANA.2018.00077

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of

the functional anatomy of language. *Cognition*, *92*(1–2), 67–99. https://doi.org/10.1016/j.cognition.2003.10.011

Hill, C. S., Coleman, M. P., & Menon, D. K. (2016). Traumatic Axonal Injury: Mechanisms and Translational Opportunities. In *Trends in Neurosciences* (Vol. 39, Issue 5, pp. 311–324). Elsevier Ltd. https://doi.org/10.1016/j.tins.2016.03.002

Hulkower, M. B., Poliak, D. B., Rosenbaum, S. B., Zimmerman, M. E., & Lipton, M. L. (2013). A Decade of DTI in Traumatic Brain Injury: 10 Years and 100 Articles Later. *AJNR Am J Neuroradiol*, *34*(11), 2064–2074. https://doi.org/10.3174/ajnr.A3395

Inglese, M., Makani, S., Johnson, G., Cohen, B. A., Silver, J. A., Gonen, O., & Grossman, R. I. (2005). Diffuse axonal injury in mild traumatic brain injury: A diffusion tensor imaging study. *Journal of Neurosurgery*, *103*(2), 298–303. https://doi.org/10.3171/jns.2005.103.2.0298

Izzy, S., Compton, R., Carandang, R., Hall, W., & Muehlschlegel, S. (2013). Self-fulfilling prophecies through withdrawal of care: do they exist in traumatic brain injury, too? *Neurocritical Care*, *19*(3), 347–363. https://doi.org/10.1007/S12028-013-9925-Z

Johnson, V. E., Stewart, W., & Smith, D. H. (2013). Axonal pathology in traumatic brain injury. In *Experimental Neurology* (Vol. 246, pp. 35–43). Exp Neurol. https://doi.org/10.1016/j.expneurol.2012.01.013

Jolly, A. E., Bălăeţ, M., Azor, A., Friedland, D., Sandrone, S., Graham, N. S. N., Zimmerman, K., & Sharp, D. J. (2020). Detecting axonal injury in individual patients after traumatic brain injury. *Brain*. https://doi.org/10.1093/brain/awaa372

Jones, E., Oliphant, T., & Peterson, P. (2001). *SciPy: Open Source Scientific Tools for Python*. http://www.scipy.org

Jones, Travis, A. R., Eden, G., Pierpaoli, C., & Basser, P. J. (2005). PASTA: Pointwise assessment of streamline tractography attributes. *Magnetic Resonance in Medicine*, *53*(6), 1462–1467. https://doi.org/10.1002/mrm.20484

Kim, N., Branch, C. A., Kim, M., & Lipton, M. L. (2013). Whole Brain Approaches for Identification of Microstructural Abnormalities in Individual Patients: Comparison of Techniques Applied to Mild Traumatic Brain Injury. *PLoS ONE*, *8*(3), 59382. https://doi.org/10.1371/journal.pone.0059382

Kinnunen, K. M., Greenwood, R., Powell, J. H., Leech, R., Hawkins, P. C., Bonnelle, V., Patel, M. C., Counsell, S. J., & Sharp, D. J. (2011). White matter damage and cognitive impairment after traumatic brain injury. *Brain*, *134*(2), 449–463. https://doi.org/10.1093/BRAIN/AWQ347

Li, J., Li, X. Y., Feng, D. F., & Gu, L. (2011). Quantitative evaluation of microscopic injury with diffusion tensor imaging in a rat model of diffuse axonal injury. *European Journal of Neuroscience*, *33*(5), 933–945. https://doi.org/10.1111/j.1460-9568.2010.07573.x

Lindemer, E. R., Salat, D. H., Smith, E. E., Nguyen, K., Fischl, B., & Greve, D. N. (2015). White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from nonconverters. *Neurobiology of Aging*, *36*(9), 2447–2457. https://doi.org/10.1016/j.neurobiolaging.2015.05.011

Mac Donald, C. L., Johnson, A. M., Cooper, D., Nelson, E. C., Werner, N. J., Shimony, J. S., Snyder, A. Z., Raichle, M. E., Witherow, J. R., Fang, R., Flaherty, S. F., & Brody, D. L. (2011). Detection of blast-related traumatic brain injury in U.S. military personnel. *The New England Journal of Medicine*, *364*(22), 2091–2100. https://doi.org/10.1056/NEJMOA1008069

Maffei, C., Lee, C., Planich, M., Ramprasad, M., Ravi, N., Trainor, D., Urban, Z., Kim, M., Jones, R. J., Henin, A., Hofmann, S. G., Pizzagalli, D. A., Auerbach, R. P., Gabrieli, J. D. E., Whitfield-Gabrieli, S., Greve, D. N., Haber, S. N., & Yendiki, A. (2021). Using diffusion MRI data acquired with ultra-high gradient strength to improve tractography in routine-quality data. *NeuroImage*, *245*, 118706. https://doi.org/10.1016/J.NEUROIMAGE.2021.118706

Maffei, C., Sarubbo, S., & Jovicich, J. (2019). A missing connection: A review of the macrostructural

anatomy and tractography of the acoustic radiation. *Frontiers in Neuroanatomy*, *13*. https://doi.org/10.3389/fnana.2019.00027

Mahalanobis, P. C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 49–55.

Makris, N., Kennedy, D. N., McInerney, S., Sorensen, A. G., Wang, R., Caviness, V. S., & Pandya, D. N. (2005). Segmentation of subcomponents within the superior longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. *Cerebral Cortex*, *15*(6), 854–869. https://doi.org/10.1093/cercor/bhh186

Makris, N., Papadimitriou, G. M., Kaiser, J. R., Sorg, S., Kennedy, D. N., & Pandya, D. N. (2009). Delineation of the middle longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. *Cerebral Cortex*, *19*(4), 777–785. https://doi.org/10.1093/cercor/bhn124

Mars, R. B., Foxley, S., Verhagen, L., Jbabdi, S., Sallet, J., Noonan, M. A. P., Neubert, F. X., Andersson, J. L., Croxson, P. L., Dunbar, R. I. M., Khrapitchev, A. A., Sibson, N. R., Miller, K. L., & Rushworth, M. F. S. (2016). The extreme capsule fiber complex in humans and macaque monkeys: a comparative diffusion MRI tractography study. *Brain Structure and Function*, *221*(8), 4059–4071. https://doi.org/10.1007/s00429-015-1146-0

McGinn, M. J., & Povlishock, J. T. (2016). Pathophysiology of Traumatic Brain Injury. *Neurosurgery Clinics*, *27*(4), 397–407. https://doi.org/10.1016/J.NEC.2016.06.002

Moe, H. K., Follestad, T., Andelic, N., Håberg, A. K., Flusund, A. M. H., Kvistad, K. A., Saksvoll, E. H., Olsen, Ø., Abel-Grüner, S., Sandrød, O., Skandsen, T., Vik, A., & Moen, K. G. (2020). Traumatic axonal injury on clinical MRI: association with the Glasgow Coma Scale score at scene of injury or at admission and prolonged posttraumatic amnesia. *Journal of Neurosurgery*, *1*(aop), 1–12. https://doi.org/10.3171/2020.6.JNS20112

Moen, K. G., Brezova, V., Skandsen, T., Håberg, A. K., Folvik, M., & Vik, A. (2014). Traumatic axonal injury: the prognostic value of lesion load in corpus callosum, brain stem, and thalamus in different magnetic resonance imaging sequences. *Journal of Neurotrauma*, *31*(17), 1486–1496. https://doi.org/10.1089/neu.2013.3258

Newcombe, V. F. J., Williams, G. B., Nortje, J., Bradley, P. G., Harding, S. G., Smielewski, P., Coles, J. P., Maiya, B., Gillard, J. H., Hutchinson, P. J., Pickard, J. D., Carpenter, T. A., & Menon, D. K. (2007). Analysis of acute traumatic axonal injury using diffusion tensor imaging. *British Journal of Neurosurgery*, *21*(4), 340–348. https://doi.org/10.1080/02688690701400882

Newcombe, V. F. J., Williams, G. B., Scoffings, D., Cross, J., Carpenter, T. A., Pickard, J. D., & Menon, D. K. (2010). Aetiological differences in neuroanatomy of the vegetative state: insights from diffusion tensor imaging and functional implications. *Journal of Neurology, Neurosurgery & Psychiatry*, *81*(5), 552–561. https://doi.org/10.1136/JNNP.2009.196246

Nolan, A. L., Petersen, C., Iacono, D., Donald, C. L. Mac, Mukherjee, P., Van Der Kouwe, A., Jain, S., Stevens, A., Diamond, B. R., Wang, R., Markowitz, A. J., Fischl, B., Perl, D. P., Manley, G. T., Keene, C. D., Diaz-Arrastia, R., & Edlow, B. L. (2020). Tractography-Pathology Correlations in Traumatic Brain Injury: A TRACK-TBI Study. *BioRxiv*, 2020.07.20.209668. https://doi.org/10.1101/2020.07.20.209668

O'Phelan, K. H., Otoshi, C. K., Ernst, T., & Chang, L. (2018). Common Patterns of Regional Brain Injury Detectable by Diffusion Tensor Imaging in Otherwise Normal-Appearing White Matter in Patients with Early Moderate to Severe Traumatic Brain Injury. *Journal of Neurotrauma*, *35*(5), 739. https://doi.org/10.1089/NEU.2016.4944

Ordóñez-Rubiano, E. G., Johnson, J., Enciso-Olivera, C. O., Marín-Muñoz, J. H., Cortes-Lozano, W., Baquero-Herrera, P. E., Ordóñez-Mora, E. G., & Cifuentes-Lobelo, H. A. (2017). Reconstruction of the Ascending Reticular Activating System with Diffusion Tensor Tractography in Patients with a Disorder of Consciousness after Traumatic Brain Injury. *Cureus*, *9*(9).

https://doi.org/10.7759/CUREUS.1723
Perlbarg, V., Puybasset, L., Tollard, E., Lehéricy, S., Benali, H., & Galanaud, D. (2009). Relation between brain lesion location and clinical outcome in patients with severe traumatic brain injury: A diffusion tensor imaging study using voxel-based approaches. *Human Brain Mapping*, *30*(12), 3924–3933. https://doi.org/10.1002/HBM.20817/FORMAT/PDF
Pieri, V., Sanvito, F., Riva, M., Petrini, A., Rancoita, P. M. V., Cirillo, S., Iadanza, A., Bello, L., Castellano, A., & Falini, A. (2021). Along-tract statistics of neurite orientation dispersion and density imaging diffusion metrics to enhance MR tractography quantitative analysis in healthy controls and in patients with brain tumors. *Human Brain Mapping*, *42*(5), 1268–1286. https://doi.org/10.1002/HBM.25291
Rheault, F., De Benedictis, A., Daducci, A., Maffei, C., Tax, C. M. W., Romascano, D., Caverzasi, E., Morency, F. C., Corrivetti, F., Pestilli, F., Girard, G., Theaud, G., Zemmoura, I., Hau, J., Glavin, K., Jordan, K. M., Pomiecko, K., Chamberland, M., Barakovic, M., … Descoteaux, M. (2020). Tractostorm: The what, why, and how of tractography dissection reproducibility. *Human Brain Mapping*, *41*(7). https://doi.org/10.1002/hbm.24917
Ripollés, P., Marco-Pallarés, J., de Diego-Balaguer, R., Miró, J., Falip, M., Juncadella, M., Rubio, F., & Rodriguez-Fornells, A. (2012). Analysis of automated methods for spatial normalization of lesioned brains. *NeuroImage*, *60*(2), 1296–1306. https://doi.org/10.1016/J.NEUROIMAGE.2012.01.094
RStudio Team. (2020). *R: A Language and Environment for Statistical Computing*.
Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 18035–18040. https://doi.org/10.1073/pnas.0805234105
Schiff, N. D. (2015). Cognitive Motor Dissociation Following Severe Brain Injuries. *JAMA Neurology*, *72*(12), 1413–1415. https://doi.org/10.1001/JAMANEUROL.2015.2899
Schilling, K. G., Rheault, F., Petit, L., Hansen, C. B., Nath, V., Yeh, F. C., Girard, G., Barakovic, M., Rafael-Patino, J., Yu, T., Fischi-Gomez, E., Pizzolato, M., Ocampo-Pineda, M., Schiavi, S., Canales-Rodríguez, E. J., Daducci, A., Granziera, C., Innocenti, G., Thiran, J. P., … Descoteaux, M. (2021). Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *NeuroImage*, *243*, 118502. https://doi.org/10.1016/J.NEUROIMAGE.2021.118502
Schnakers, C., Bessou, H., Rubi-Fessen, I., Hartmann, A., Fink, G. R., Meister, I., Giacino, J. T., Laureys, S., & Majerus, S. (2015). Impact of aphasia on consciousness assessment: a cross-sectional study. *Neurorehabilitation and Neural Repair*, *29*(1), 41–47. https://doi.org/10.1177/1545968314528067
Schweitzer, A. D., Niogi, S. N., Whitlow, C. T., & Tsiouris, A. J. (2019). Traumatic brain injury: Imaging patterns and complications. *Radiographics*, *39*(6), 1571–1595. https://doi.org/10.1148/RG.2019190076/ASSET/IMAGES/LARGE/RG.2019190076.FIG11F.JPEG
Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.
Setsompop, K., Cohen-Adad, J., Gagoski, B. A., Raij, T., Yendiki, A., Keil, B., Wedeen, V. J., & Wald, L. L. (2012). Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *NeuroImage*, *63*(1), 569–580. https://doi.org/10.1016/j.neuroimage.2012.06.033
Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., Watkins, K. E., Ciccarelli, O., Cader, M. Z., Matthews, P. M., & Behrens, T. E. J. (2006). Tract-based spatial statistics: Voxelwise analysis of multi-subject diffusion data. *NeuroImage*, *31*(4), 1487–1505. https://doi.org/10.1016/J.NEUROIMAGE.2006.02.024
Snider, S. B., Bodien, Y. G., Bianciardi, M., Brown, E. N., Wu, O., & Edlow, B. L. (2019). Disruption of the ascending arousal network in acute traumatic disorders of consciousness. *Neurology*, *93*(13),

E1281–E1287. https://doi.org/10.1212/WNL.0000000000008163
Tamashiro, M., Cozzo, D., Mattei, M., Salierno, F., Rivas, M. E., Alzúa, O., Olmos, L., Bonamico, L., & Leiguarda, R. (2012). Early motor predictors of recovery in patients with severe traumatic brain injury. *Https://Doi.Org/10.3109/02699052.2012.661911*, *26*(7–8), 921–926. https://doi.org/10.3109/02699052.2012.661911
Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness. A practical scale. *Lancet (London, England)*, *2*(7872), 81–84. https://doi.org/10.1016/S0140-6736(74)91639-0
Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., & Murray, G. (2014). The Glasgow Coma Scale at 40 years: standing the test of time. *The Lancet Neurology*, *13*(8), 844–854. https://doi.org/10.1016/S1474-4422(14)70120-6
Thibaut, A., Bodien, Y. G., Laureys, S., & Giacino, J. T. (2020). Minimally conscious state "plus": diagnostic criteria and relation to functional recovery. *Journal of Neurology*, *267*(5), 1245–1254. https://doi.org/10.1007/S00415-019-09628-Y
Turgeon, A. F., Lauzier, F., Simard, J. F., Scales, D. C., Burns, K. E. A., Moore, L., Zygun, D. A., Bernard, F., Meade, M. O., Dung, T. C., Ratnapalan, M., Todd, S., Harlock, J., & Fergusson, D. A. (2011). Mortality associated with withdrawal of life-sustaining therapy for patients with severe traumatic brain injury: a Canadian multicentre cohort study. *CMAJ : Canadian Medical Association Journal = Journal de l'Association Medicale Canadienne*, *183*(14), 1581–1588. https://doi.org/10.1503/CMAJ.101786
Ubukata, S., Ueda, K., Sugihara, G., Yassin, W., Aso, T., Fukuyama, H., & Murai, T. (2016). Corpus callosum pathology as a potential surrogate marker of cognitive impairment in diffuse axonal injury. *Journal of Neuropsychiatry and Clinical Neurosciences*, *28*(2), 97–103. https://doi.org/10.1176/APPI.NEUROPSYCH.15070159/ASSET/IMAGES/LARGE/APPI.NEUROPSYCH.15070159F3.JPEG
Van Hecke, W., Leemans, A., & Emsell, L. (2016). DTI Analysis Methods: Voxel-Based Analysis. *Diffusion Tensor Imaging: A Practical Handbook*, 183–203. https://doi.org/10.1007/978-1-4939-3118-7_10
Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*. https://doi.org/10.1007/978-0-387-21706-2
Wang, J. Y., Bakhadirov, K., Devous, M. D., Abdi, H., McColl, R., Moore, C., Marquez De La Plata, C. D., Ding, K., Whittemore, A., Babcock, E., Rickbeil, T., Dobervich, J., Kroll, D., Dao, B., Mohindra, N., Madden, C. J., & Diaz-Arrastia, R. (2008). Diffusion Tensor Tractography of Traumatic Diffuse Axonal Injury. *Archives of Neurology*, *65*(5), 619–626. https://doi.org/10.1001/ARCHNEUR.65.5.619
Warner, M. A., De La Plata, C. M., Spence, J., Wang, J. Y., Harper, C., Moore, C., Devous, M., & Diaz-Arrastia, R. (2010). Assessing Spatial Relationships between Axonal Integrity, Regional Brain Volumes, and Neuropsychological Outcomes after Traumatic Axonal Injury. *Journal of Neurotrauma*, *27*(12), 2121. https://doi.org/10.1089/NEU.2010.1429
Welniarz, Q., Dusart, I., & Roze, E. (2017). The corticospinal tract: Evolution, development, and human disorders. *Developmental Neurobiology*, *77*(7), 810–829. https://doi.org/10.1002/DNEU.22455
Xu, J., Rasmussen, I. A., Lagopoulos, J., & Haberg, A. (2007). Diffuse axonal injury in severe traumatic brain injury visualized using high-resolution diffusion tensor imaging. *Journal of Neurotrauma*, *24*(5), 753–765. https://doi.org/10.1089/neu.2006.0208
Yendiki, A., Koldewyn, K., Kakunoori, S., Kanwisher, N., & Fischl, B. (2014). Spurious group differences due to head motion in a diffusion MRI study. *NeuroImage*, *88*, 79–90. https://doi.org/10.1016/j.neuroimage.2013.11.027
Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., Wang, R., Salat, D., Ehrlich, S., Behrens, T., Jbabdi, S., Gollub, R., & Fischl, B. (2011). Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. *Frontiers in Neuroinformatics*, *5*, 23. https://doi.org/10.3389/fninf.2011.00023

Yendiki, A., Reuter, M., Wilkens, P., Rosas, H. D., & Fischl, B. (2016). Joint reconstruction of white-matter pathways from longitudinal diffusion MRI data with anatomical priors. *NeuroImage*, *127*, 277–286. https://doi.org/10.1016/j.neuroimage.2015.12.003

Yuh, E. L., Cooper, S. R., Mukherjee, P., Yue, J. K., Lingsma, H. F., Gordon, W. A., Valadka, A. B., Okonkwo, D. O., Schnyer, D. M., Vassar, M. J., Maas, A. I. R., Manley, G. T., Casey, S. S., Cheong, M., Dams-O'Connor, K., Hricik, A. J., Inoue, T., Menon, D. K., Morabito, D. J., … Sinha, T. K. (2014). Diffusion Tensor Imaging for Outcome Prediction in Mild Traumatic Brain Injury: A TRACK-TBI Study. *Https://Home.Liebertpub.Com/Neu*, *31*(17), 1457–1477. https://doi.org/10.1089/NEU.2013.3171

Zhang, J., Wei, R. L., Peng, G. P., Zhou, J. J., Wu, M., He, F. P., Pan, G., Gao, J., & Luo, B. Y. (2017). Correlations between diffusion tensor imaging and levels of consciousness in patients with traumatic brain injury: a systematic review and meta-analysis. *Scientific Reports 2017 7:1*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-02950-3

Zöllei, L., Jaimes, C., Saliba, E., Grant, P. E., & Yendiki, A. (2019). TRActs constrained by UnderLying INfant anatomy (TRACULInA): An automated probabilistic tractography tool with anatomical priors for use in the newborn brain. *NeuroImage*, *199*, 1–17. https://doi.org/10.1016/j.neuroimage.2019.05.051