

TractSpLearn: Specialized Shared-Manifold Learning for Individualized Detection of Subtle White Matter Alterations in Mild Traumatic Brain Injury

Jiqing Huang ¹, Ali Al-Husseini ^{2#}, Yi Chen ^{3,4#}, Anna Gard ²,
Laurent Lamalle ¹, Mohamed Ali Bahri ¹, Markus Nilsson ⁵,
Niklas Marklund ^{2†}, Christophe Phillips ^{1†}, Evgenios N. Kornaropoulos ^{1,5,6*†}

¹ GIGA – CRC Human Imaging, University of Liège, Belgium

² Department of Clinical Sciences Lund, Neurosurgery, Lund university, Skane University Hospital, Sweden

³ AMIIC Lab, Guizhou University, China

⁴ D-Lab, Maastricht University, Netherlands

⁵ Department of Clinical Sciences Lund, Diagnostic Radiology, Medical Faculty, Lund University, Sweden

⁶ Aix-Marseille Univ, CNRS, CRMBM, Marseille, France

* Correspondence: ekornaropoulos@uliege.be

Equal contribution as second authors † Equal contribution as last authors

September 12, 2026

Keywords: Diffusion MRI, Diffusion Kurtosis Imaging, Fiber Tractography, Manifold Learning, Regression, Traumatic Brain Injury

Abstract:

Traumatic brain injury (TBI) often leads to subtle white matter damage that remains undetected on conventional MRI. Diffusion kurtosis imaging (DKI), an extension of diffusion tensor imaging (DTI), provides complementary information on non-Gaussian water diffusion and is sensitive to complex white-matter microstructure. With the advent of ultra-high-field MRI, the spatial resolution and signal-to-noise ratios (SNR) have been significantly enhanced, enabling more precise visualization of subtle abnormalities. Building on these advances, we developed TractSpLearn, an individualized tract-based learning framework that jointly considers within-group variability and between-group differences. Unlike the original TractLearn framework, which learns a normative manifold exclusively from healthy controls, TractSpLearn incorporates both healthy controls and patients to learn a shared manifold with a healthy-anchored representation and an additional patient-related component. To assess the performance of the proposed method, we compared TractSpLearn with the original TractLearn in three cohorts: (i) healthy controls (HC), (ii) athletes with persistent post-concussive syndromes (PPCS), and (iii) athletes with repeated head injuries (RHI), with abnormalities particularly evident in axial kurtosis (AK) and mean diffusivity (MD). In RHI, TractSpLearn highlighted recurrent abnormalities across patients. In the PPCS cohort, the overall group-level differences were more modest, potentially reflecting both limited statistical power due to the small sample size and partial normalization of white-matter alterations during recovery. Still TractSpLearn identified abnormality evidence in more patients and across more affected tracts than TractLearn.

1 Introduction

Traumatic brain injury (TBI) refers to brain damage caused by an external mechanical force, such as a forceful bump, blow, or jolt to the head, or skull penetration by an object. Beyond the initial impact, TBI often triggers a cascade of secondary pathophysiological processes. Patients with TBI may experience persistent cognitive impairments, emotional and behavioral disturbances, and sensory or motor deficits. In severe cases, these consequences can lead to long-term disability, with outcomes varying according to the extent and, particularly, the location of the injury.

Magnetic resonance imaging (MRI) has been extensively used in TBI to investigate macroscopic abnormalities, such as contusions, hemorrhage, and edema. However, non-hemorrhagic lesions, including diffuse axonal injury, often remain undetectable on conventional MRI, limiting its ability to capture subtle microstructural damage that may underlie persistent symptoms (Lee, 2020). To address this limitation, diffusion tensor imaging (DTI) provides more sensitive and quantitative information by characterizing white matter microstructural integrity. For example, it can reveal alterations associated with axonal organization, demyelination, and connectivity disruption by mapping water diffusion along white matter tracts. Building upon DTI, more advanced diffusion models, such as diffusion kurtosis imaging (DKI), further probe tissue heterogeneity and microstructural complexity. Compared to DTI, DKI is sensitive to microstructural complexity potentially associated with axonal beading, cytotoxic edema and microglial activation, all of which are common in TBI. In practice, DKI requires higher b-values, a larger number of diffusion-encoding directions, and high image quality to reliably capture non-Gaussian behavior. This poses a challenge for MRI acquisition, especially for clinical settings, where reproducibility and stability are limited. Indeed, Mahan et al. (2021) noted that increases, decreases, and no changes in DTI metrics have been reported in TBI studies. Huang et al. (2022) explained these inconsistencies by highlighting the roles of injury severity, injury type, time since injury, and sample characteristics. Through longitudinal monitoring of early and follow-up examinations, Muftuler et al. (2020) demonstrated that DKI was more sensitive than DTI and notably revealed delayed microstructural alterations emerging after injury. Although our understanding of the full spectrum of brain tissue injury and recovery processes remains limited, there is a general consensus that higher spatial resolution and signal-to-noise ratio (SNR) improve the ability to delineate fine-scale white matter architecture and reduce partial volume effects (Kornaropoulos et al., 2022). Choi et al. (2011) demonstrated that 7T provides a substantially higher SNR, resulting in more stable and less variable quantitative measurements. This technical advantage creates the opportunity to detect more subtle microstructural abnormalities. Conventional strategies for analyzing DTI, such as GLM-based voxel-wise analyses and tract-based spatial statistics, are primarily designed for group studies Genç et al. (2025) and Hashim et al. (2017). Their focus on population-level

inference and sensitivity to cohort heterogeneity make them less suited to detecting subtle abnormalities at the individual level.

To overcome these limitations, Attyé et al. (2021) proposed TractLearn, which learns tract-specific latent representations from healthy controls and back-projects subject-level reconstruction deviations onto diffusion-parametric images. Building on this approach, Roger et al. (2022) applied TractLearn in patients with chronic temporal lobe epilepsy, and Clemente et al. (2023) used TractLearn to profile moderate to severe TBI, demonstrating that TractLearn can generate individualized white matter signatures and reveal inter-patient heterogeneity that remains hidden in conventional group-based analyses. However, the original TractLearn framework models individual deviations relative to healthy variability but does not explicitly incorporate patient-group information or optimize patient–control separability. As a result, although TractLearn may be sensitive to individual deviations from the healthy manifold, these deviations are not necessarily specific to patient-related microstructural alterations. Taken together, these studies reveal a key gap: although advanced diffusion models and 7T MRI improve sensitivity to microstructural damage, current individualized analytical frameworks remain challenged by the balance between sensitivity to subject-specific deviations and specificity for patient-related alterations, particularly in studies with limited sample sizes. This raises the question of whether detected abnormalities reflect true pathology or normal variation. Addressing this limitation is essential for translating diffusion-MRI findings into clinically meaningful patient-level assessments. In this context, we investigate DKI-derived white-matter alterations in athletes with persistent post-concussive symptoms (PPCS) and athletes exposed to repetitive head impacts (RHI), using state-of-the-art 7T DKI. To improve the balance between sensitivity to individualized abnormalities and specificity for patient-related alterations, we introduce TractSpLearn, an extension of TractLearn that incorporates both between-group differences and within-group variability into the learning framework. We compared TractSpLearn with the original TractLearn across multiple white-matter tracts and diffusion-derived parameters to evaluate whether integrating individual deviations and patient-group information improves the detection of patient-related microstructural alterations in limited-sample settings.

2 Materials and Methods

2.1 Participants

This retrospective study evaluated the performance of TractLearn and TractSpLearn on a cohort of fifty-six participants. Ethical approval was obtained from the institutional review board of Lund University,

Sweden (Dnr 2017/1049), and all procedures complied with the Declaration of Helsinki. Written informed consent was obtained from all participants. The cohort consisted of 20 healthy controls (HC), 16 athletes with persistent post-concussive syndromes (PPCS), and 20 athletes with a history of repeated head injuries (RHI).

The HC group comprised individuals without a history of competitive contact sports who were age- and sex-matched to the athletes, each engaging in ≥ 3 exercise sessions per week for the preceding five years. The PPCS group included former elite athletes who continued to experience post-concussive symptoms for at least six months after concussion, with a severity sufficient to prevent return to sport and to limit work, school, or regular social activities. The RHI group consisted of active athletes with a documented history of recurrent head injuries and ongoing exposure to repetitive head impacts during sport.

2.2 MRI acquisition

Both structural and diffusion-weighted imaging (DWI) were acquired on a 7T Philips Achieva scanner. High-resolution anatomical images were obtained using a 3D T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) sequence with the following parameters: field of view (FOV): $230 \times 230 \times 180 \text{ mm}^3$, resolution: $0.80 \times 0.80 \times 0.80 \text{ mm}^3$, repetition time (TR): 8.00 ms, echo time (TE): 1.97 ms. DWI data were collected with FOV = $224 \times 224 \times 110 \text{ mm}^3$, resolution = $2 \times 2 \times 2 \text{ mm}^3$, TR = 9200 ms, TE = 65 ms, using multiple diffusion-encoding directions: two $b = 0 \text{ s/mm}^2$ volumes with reversed phase-encoding directions for distortion correction, 6 directions at $b = 100 \text{ s/mm}^2$, 6 directions at $b = 500 \text{ s/mm}^2$, 10 directions at $b = 1000 \text{ s/mm}^2$, and 30 directions at $b = 2000 \text{ s/mm}^2$.

2.3 DWI preprocessing and DKI fitting

DWI data were preprocessed using a standardized pipeline to ensure data quality and consistency across participants (Gard et al., 2024). The pipeline included: (i) denoising using Marchenko–Pastur principal component analysis (Veraart et al., 2016); (ii) Gibbs ringing correction using a non-iterative approach (Kellner et al., 2016); (iii) TOPUP susceptibility distortion correction using a reversed diffusion encoding $b=0$ acquisition, and (iv) eddy current and motion correction. After preprocessing, diffusion and kurtosis metrics, including fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD), radial diffusivity (RD), mean kurtosis (MK), axial kurtosis (AK), and radial kurtosis (RK), were estimated using weighted linear least squares as implemented in DIPY (DiffusionKurtosisModel) (Garyfallidis et al., 2014).

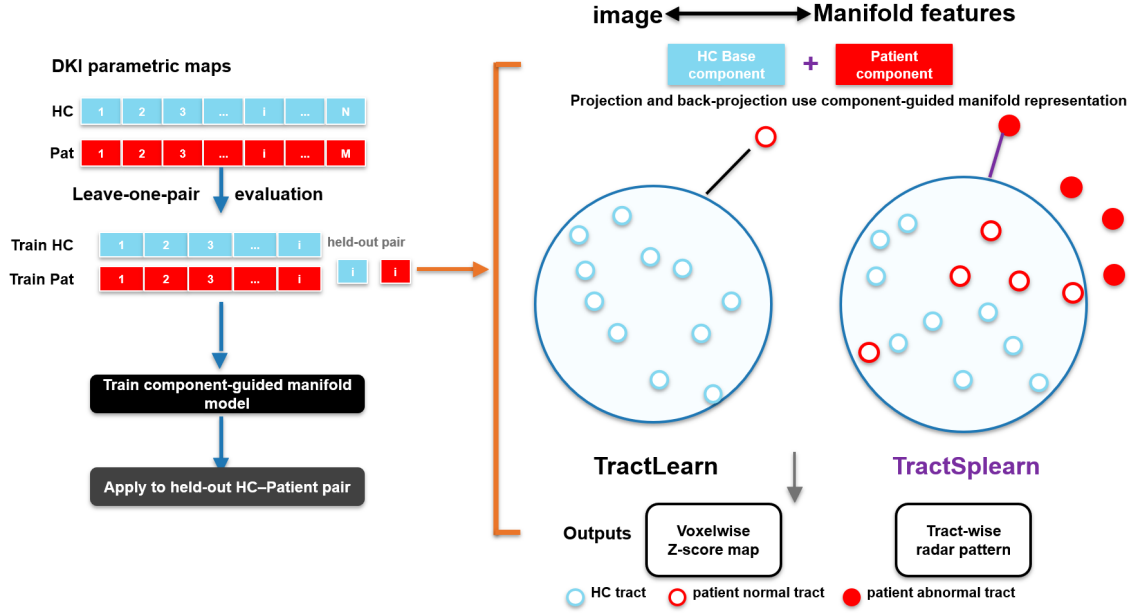


Figure 1: Flowchart of TractLearn and of TractSpLearn and key differences between TractLearn and TractSpLearn

2.4 TractSpLearn : TractLearn based from work with individual and group criterion

The proposed method, referred to as TractSpLearn, builds upon the original TractLearn framework developed by (Attyé et al., 2021). TractLearn and TractSpLearn were applied independently to each white-matter tract and each diffusion-derived parameter. White-matter tracts were segmented using TractSeg resulting in 72 tract-specific segmentations for each DKI parameter (Wasserthal et al., 2018).

The patient (P) cohorts, PPCS and RHI, were analyzed separately using the same HC cohort as the reference group. For each cohort, we implemented a paired leave-one-out (LOO) evaluation strategy, as illustrated in figure 1. In each fold, one HC-PAT pair, matched by subject index, was reserved for testing. For PPCS, 16 HC-patient pairs were evaluated, with the remaining four HCs retained in each training fold. For RHI, all 20 HC-patient pairs were evaluated. The same held-out pair was evaluated using both TractLearn and TractSpLearn. TractLearn was trained using only the remaining HC, whereas TractSpLearn was trained using the remaining HC together with the remaining patients. For each LOO training fold, TractLearn and TractSpLearn were applied independently to each white-matter tract and each diffusion-derived parameter. The processing pipeline consisted of manifold feature extraction, forward mapping, back-projection, and healthy-reference z-score calculation.

Features extraction Unlike the original TractLearn framework, in which the manifold features were derived exclusively from HCs, TractSpLearn extracted shared manifold features from both HCs and patients in the training set. For each subject i and white-matter tract j , a set of manifold features $f_{i,j}$ was extracted from the high-dimensional DKI parameter vector $p_{i,j}$, where

$$p \in P = \{FA, MD, AD, RD, MK, AK, RK\}. \quad (1)$$

The manifold features were defined as

$$f_{i,j} = U(p_{i,j}), \quad (2)$$

where $U(\cdot)$ denotes the UMAP manifold-learning method (McInnes et al., 2018). UMAP projected the HC and patient training data into a shared latent manifold space. The resulting representation captures the common tract structure and inter-subject variability present in both groups.

Forward mapping (projection to the manifold) To enable inference on unseen subjects, a forward nonlinear regression model was trained to estimate the manifold coordinates directly from $p_{i,j}$:

$$\hat{f}_{i,j} = \Psi(p_{i,j}). \quad (3)$$

The forward mapping was parameterized as a healthy-anchored base component together with an additional patient-related component:

$$\Psi(p_{i,j}) = \Psi_{\text{base}}(p_{i,j}) + y_i \Psi_{\text{component}}(p_{i,j}) + \epsilon_{i,j}, \quad (4)$$

where $y_i \in \{0, 1\}$ indicates the group label, with $y_i = 0$ for HCs and $y_i = 1$ for patients. Here, $\Psi_{\text{base}}(\cdot)$ represents the healthy-anchored base component, $\Psi_{\text{component}}(\cdot)$ represents the additional patient-related component, and $\epsilon_{i,j}$ is the residual term. For each held-out subject, two latent representations were generated independently of the subject's group label. The forced HC representation was defined as

$$\hat{f}_{i,j}^{\text{HC}} = \Psi_{\text{base}}(p_{i,j}), \quad (5)$$

whereas the forced patient representation was defined as

$$\hat{f}_{i,j}^{\text{PAT}} = \Psi_{\text{base}}(p_{i,j}) + \Psi_{\text{component}}(p_{i,j}). \quad (6)$$

Back-projection (reconstruction in parameter space) The estimated HC and patient manifold representations were reconstructed back into the original DKI parameter space using a shared inverse nonlinear regression model $\Phi(\cdot)$:

$$\hat{p}_{i,j}^{\text{HC}} = \Phi\left(\hat{f}_{i,j}^{\text{HC}}\right), \quad (7)$$

and

$$\hat{p}_{i,j}^{\text{PAT}} = \Phi\left(\hat{f}_{i,j}^{\text{PAT}}\right). \quad (8)$$

Therefore, the difference between the two reconstructed tract vectors originated from their latent representations rather than from separate group-specific back-projection models. The corresponding reconstruction residuals were defined as

$$r_{i,j}^{\text{HC}} = p_{i,j} - \hat{p}_{i,j}^{\text{HC}}, \quad (9)$$

and

$$r_{i,j}^{\text{PAT}} = p_{i,j} - \hat{p}_{i,j}^{\text{PAT}}. \quad (10)$$

Z-score calculation Rather than directly subtracting the two reconstructed tract vectors, TractSpLearn compared the absolute reconstruction residuals obtained from the HC and patient branches. For subject i , tract j , and voxel x , the signed reconstruction-preference evidence was calculated as

$$S_{i,j}(x) = |r_{i,j}^{\text{HC}}(x)| - |r_{i,j}^{\text{PAT}}(x)|. \quad (11)$$

Equivalently,

$$S_{i,j}(x) = |p_{i,j}(x) - \hat{p}_{i,j}^{\text{HC}}(x)| - |p_{i,j}(x) - \hat{p}_{i,j}^{\text{PAT}}(x)|. \quad (12)$$

A positive value indicates that the patient branch produced a smaller reconstruction error than the HC branch, whereas a negative value indicates that the HC branch provided the better reconstruction. A voxel-wise normative reference was estimated from the signed evidence maps of the training HC using a leave-one-out procedure. The mean and standard deviation of the HC leave-one-out evidence were calculated as

$$\mu_{j,p}(x) = \text{Mean}\left(S_{h,j}^{\text{LOO}}(x)\right), \quad h \in \text{HC}_{\text{train}}, \quad (13)$$

and

$$\sigma_{j,p}(x) = \text{SD}\left(S_{h,j}^{\text{LOO}}(x)\right), \quad h \in \text{HC}_{\text{train}}. \quad (14)$$

For each held-out subject, the corresponding TractSpLearn Z-score was calculated as

$$Z_{i,j}(x) = \frac{S_{i,j}(x) - \mu_{j,p}(x)}{\sigma_{j,p}(x)}, \quad i \in \text{test}. \quad (15)$$

2.5 Evaluation metrics

As described in Fig. 1, TractLearn and TractSpLearn were evaluated at group and individual levels. The global analysis quantified the overall abnormality burden for each diffusion parameters (and further for each tracts). The individual-level analysis characterized subject-specific abnormality patterns.

Group level For each held-out subject, voxel-wise Z-score maps were obtained independently for each white-matter tract and diffusion parameter. Tracts containing fewer than 50 valid voxels in the original tract support were excluded from the quantitative analysis. In addition, voxels with extreme standardized values ($|Z| > 10$) were excluded before quantitative averaging to reduce the influence of unstable Z-scores potentially arising from very small normative variance. The tract-size criterion was determined before this voxel-level exclusion. For subject i , tract j , diffusion parameter $p \in \mathcal{P}$, and voxel $x \in V_{i,j}^*$, where $V_{i,j}^*$ denotes the retained voxels after quality-control filtering, the continuous abnormality evidence was defined as

$$S_{i,j,p}(x) = \begin{cases} |Z_{i,j,p}(x)|, & \text{TractLearn,} \\ \max(Z_{i,j,p}(x), 0), & \text{TractSpLearn.} \end{cases} \quad (16)$$

For TractLearn, both positive and negative deviations from the HC normative reference contributed to the abnormality evidence, whereas for TractSpLearn, only positive deviations indicating increased preference for the patient-related reconstruction were retained. For each subject and diffusion parameter, the global abnormality evidence was calculated by averaging the continuous evidence across all retained voxels from all eligible white-matter tracts. Thus, the global measure was voxel-weighted, with each retained voxel contributing equally to the final score. HC and patient distributions were compared independently for the PPCS and RHI cohorts using two-sided Mann–Whitney U tests with Benjamini–Hochberg false discovery rate for each DKI parameter. Individual subject values together with group means and standard deviations were reported.

Individual-level abnormality patterns To characterize subject-specific white-matter abnormalities, voxel-wise Z-score maps were converted into binary abnormality maps using method-specific criteria:

$$B_{i,j,p}(x) = \begin{cases} \mathbb{I}(|Z_{i,j,p}(x)| > T), & \text{TractLearn,} \\ \mathbb{I}(Z_{i,j,p}(x) > T), & \text{TractSpLearn,} \end{cases} \quad (17)$$

where $T = 4.25$ denotes the predefined voxel-wise Z-score threshold. For TractLearn, both positive and negative deviations from the healthy reference were considered abnormal, whereas for TractSpLearn only

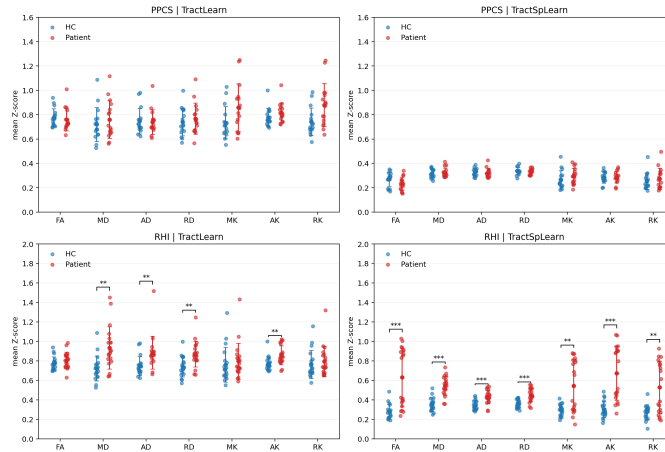


Figure 2: Group-averaged Z-score profiles for HC and patient groups

positive Z-scores were retained because positive values represent increased preference for the patient-related representation relative to the healthy reference. For each subject and diffusion parameter, a tract-wise abnormality percentage was calculated as the proportion of supra-threshold voxels among all retained voxels within each white-matter tract. This percentage was used to summarize the extent of subject-specific tract involvement. Tracts containing fewer than 50 retained voxels were excluded from the tract-level analysis to reduce instability associated with very small tract representations.

3 Result

3.1 Group-level comparison of abnormality scores

To compare the abnormality patterns identified by TractLearn and TractSpLearn, we first summarized the subject-level abnormality evidence at the group level. Of the 72 tracts, 29 containing fewer than 50 valid voxels were excluded from subsequent analyses. For each diffusion parameter, voxel-wise evidence was aggregated across the analyzed white matter tracts to obtain a global abnormality score for each subject. Figure 2 shows the group-level distributions for FA, MD, AD, RD, MK, AK, and RK, with the corresponding statistical significance indicated in the figure. Group-level mean Z-score maps for all diffusion parameters and both methods are provided in the Supplementary Material. Overall, the pattern of within-cohort HC–patient differences varied between the PPCS and RHI cohorts. In the PPCS cohort, substantial overlap between HC and patients was observed for most diffusion parameters with both methods, indicating relatively weak group-level differences. In contrast, better delineated group differences were observed in the RHI cohort. TractLearn showed significant differences for MD, AD,

RD, and AK, whereas TractSpLearn showed significant HC–RHI differences across all seven diffusion parameters, including FA, MD, AD, RD, MK, AK, and RK. Based on the group-level findings, MD and AK were selected for subsequent individual-level visualization and analysis. MD was chosen as a representative conventional diffusion metric, whereas AK showed particularly prominent abnormalities and group separation. In the following analyses, we therefore focused on these two parameters to illustrate subject-specific abnormality patterns and spatial distributions.

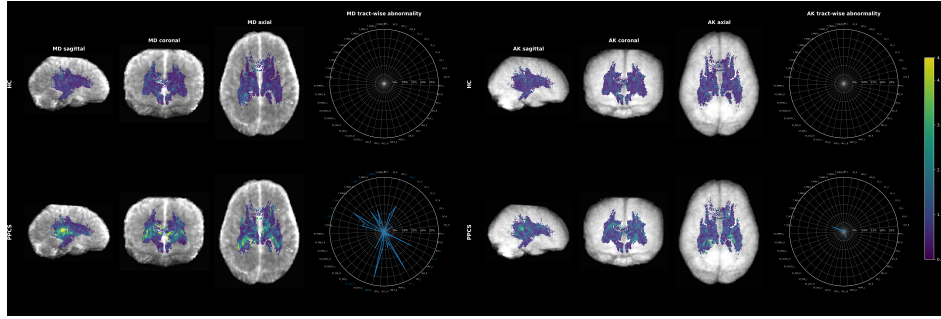
At the individual level, patients generally exhibited more extensive and spatially clustered abnormality patterns than healthy controls, although the magnitude and distribution of these alterations varied across subjects and diffusion parameters. TractLearn also produced visible supra-threshold abnormalities in healthy controls, whereas TractSpLearn generally yielded sparser patient-related evidence in controls while retaining localized abnormalities in patients. This contrast was particularly apparent in the RHI data, where TractSpLearn showed more spatially concentrated patient-related alterations with comparatively limited background evidence in healthy controls. For each representative patient and healthy control, the corresponding radar plots summarize the proportion of supra-threshold voxels within each white matter tract, providing an individual-level profile of tract involvement.

Table 1: Subject-level abnormality detection across MD and AK. A tract was considered affected when at least 15% of its retained voxels were supra-threshold in either parameter. Each affected tract was counted only once across parameters. Mean and SD of affected tract counts were calculated only among abnormality-positive subjects..

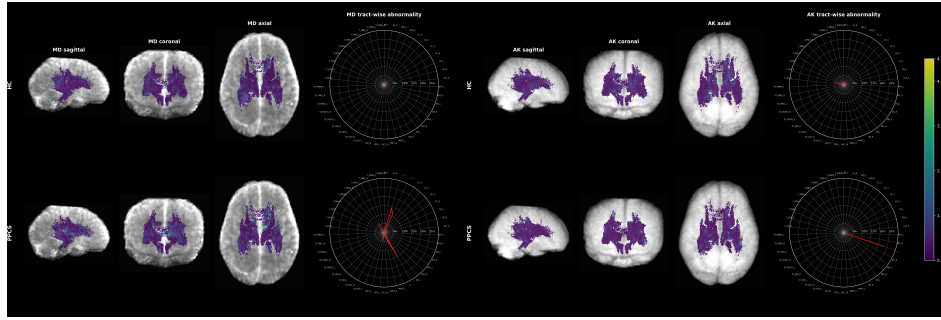
Cohort	Method	Group	Lesion-positive n/N	Affected Tracts(Total: 43) mean $\pm$ SD
PPCS	TractLearn	HC	3/16	2.33 $\pm$ 0.58
PPCS	TractLearn	Patient	2/16	4.50 $\pm$ 0.71
PPCS	TractSpLearn	HC	3/16	1.67 $\pm$ 1.15
PPCS	TractSpLearn	Patient	9/16	1.56 $\pm$ 0.53
RHI	TractLearn	HC	3/20	2.33 $\pm$ 0.58
RHI	TractLearn	Patient	7/20	2.00 $\pm$ 1.83
RHI	TractSpLearn	HC	11/20	1.27 $\pm$ 0.47
RHI	TractSpLearn	Patient	18/20	6.61 $\pm$ 4.23

To further evaluate lesion detection at the subject level, we summarized the number of participants with at least one affected white-matter tract and the number of affected tracts per subject across the selected diffusion parameters. A tract was considered affected when at least 15% of its voxels showed abnormal values according to the predefined Z-score criterion. For the combined analysis across diffusion parameters, the same tract was counted only once when it was identified as abnormal in more than one parameter. Subjects with at least one affected tract were classified as lesion-positive.

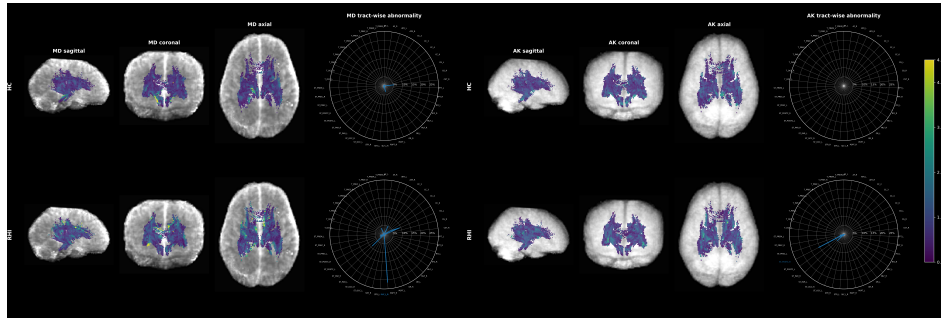
As shown in Table 1, TractSpLearn increased patient-level lesion detection in both cohorts. The mean and



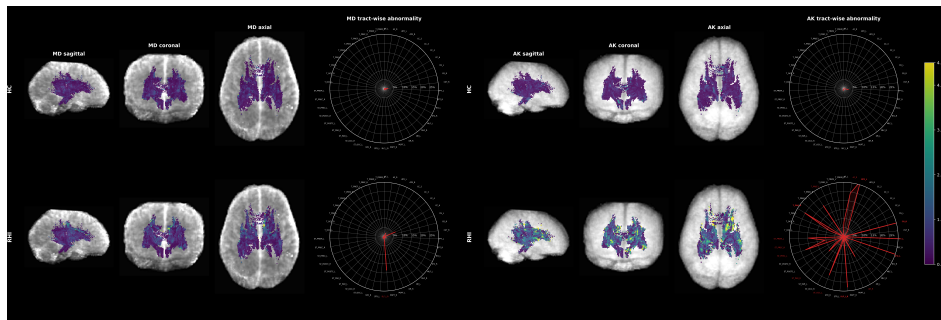
(a)



(b)



(c)

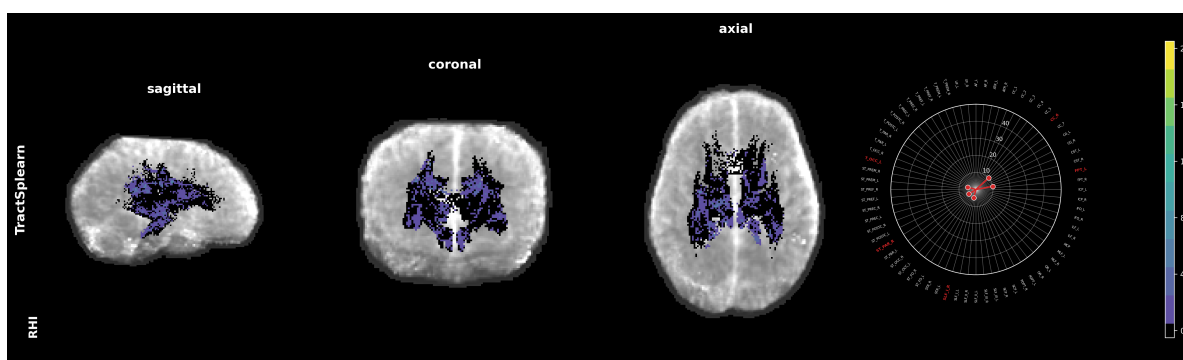


(d)

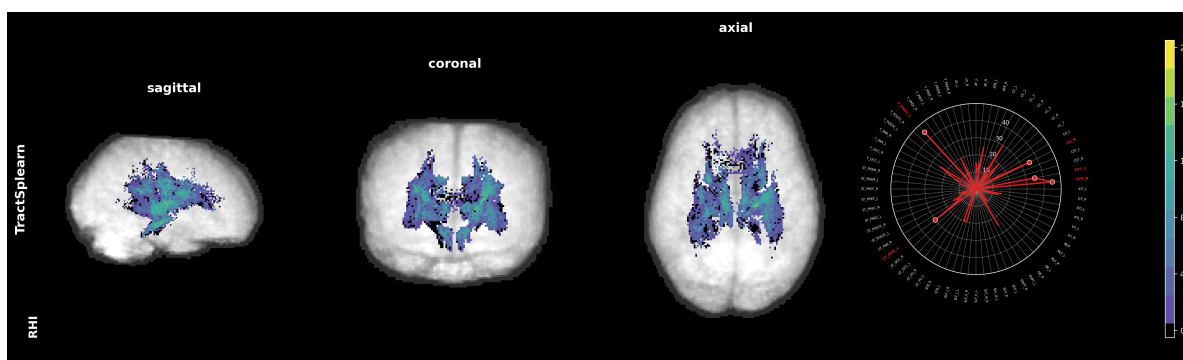
Figure 3: Individual-level tract-wise abnormality patterns in representative subjects. (a) PPCS using TractLearn; (b) PPCS using TractSpLearn; (c) RHI using TractLearn; and (d) RHI using TractSpLearn. For each sub-figures, the left half of the figure shows MD, whereas the right half shows AK. Within each parameters block, the panels from left to right show sagittal, coronal and axial views of the spatial abnormality maps, followed by the tract-wise abnormality profile shown in radar plot. The upper row shows the HC group and the lower row shows the patient group.

SD of affected tract counts were calculated only among abnormality-positive subjects. In PPCS, patient positivity increased from 2/16 (12.5%) with TractLearn to 9/16 (56.3%) with TractSpLearn, while the mean affected tract count decreased from 4.50 ± 0.71 to 1.56 ± 0.53 . HC positivity remained unchanged at 3/16 (18.8%), with the corresponding tract count decreasing from 2.33 ± 0.58 to 1.67 ± 1.15 .

In RHI, patient positivity increased from 7/20 (35.0%) to 18/20 (90.0%), accompanied by an increase in the mean affected tract count from 2.00 ± 1.83 to 6.61 ± 4.23 . HC positivity also increased, from 3/20 (15.0%) to 11/20 (55.0%), although the corresponding tract count decreased from 2.33 ± 0.58 to 1.27 ± 0.47 . Overall, these findings indicate that TractSpLearn identified abnormalities in a larger subset of patients, while in RHI, abnormality-positive patients showed greater tract involvement than abnormality-positive healthy controls.



(a)



(b)

Figure 4: Cross-subject spatial consistency and tract-wise distribution of white matter abnormalities in the RHI cohort. (a) Cross-subject abnormality distribution for MD; (b) cross-subject abnormality distribution for AK. For each sub-figures, from left to right correspond to sagittal, coronal and axial views, followed by the average cross subject abnormality radar plot

3.2 Exploratory analysis of recurrent abnormality patterns in RHI

To further explore whether TBI was associated with a reproducible pattern of white matter abnormalities across individuals, we performed an exploratory group-level analysis of the spatial and tract-wise distribution of detected abnormalities. Given the PPCS cohort showed relatively few detected abnormalities and the RHI cohort showed more extensive abnormalities, we therefore restricted the following exploratory analysis to the RHI cohort to investigate whether abnormalities recurred in similar regions across participants. Figure 4 shows the cross-subject distribution of abnormalities in the RHI cohort for MD and AK. The spatial frequency maps show the number of participants exhibiting an abnormality at each voxel, thereby highlighting regions in which abnormalities were repeatedly detected across individuals.

In parallel, the tract-wise analysis summarizes the frequency with which each white matter tract was marked as abnormal across the cohort, allowing identification of tracts that were consistently affected across participants. We observed that AK abnormalities showed a broad and spatially consistent distribution across participants, particularly involving bilateral deep and periventricular white matter regions. MD abnormalities also demonstrated recurrent spatial patterns across individuals, with several regions showing repeated involvement. The tract-wise analysis further identified a subset of white matter tracts that were frequently affected across participants for both AK and MD, with the five most frequently involved tracts highlighted in red.

4 Discussion

In this study, we proposed TractSpLearn, a tract-wise framework for individualized characterization of white matter abnormalities by explicitly incorporating patient-related information into the modeling process. Unlike the original TractLearn framework, which identifies individual abnormalities primarily according to deviations from a healthy reference, TractSpLearn models both shared structural characteristics and an additional patient-related component. For an unseen subject, the resulting abnormality measure therefore reflects the relative preference between the healthy and patient representations rather than deviation from the healthy population alone. This formulation provides a more disease-informed characterization of individualized diffusion abnormalities while retaining healthy controls as the normative reference.

The main finding was that TractSpLearn provided spatially refined group-level differentiation between healthy controls and patients, particularly in the RHI cohort. Compared with TractLearn, the proposed framework generally maintained relatively low abnormality evidence in healthy controls while preserving

or enhancing patient-related deviations across multiple diffusion parameters. These results suggest that explicitly modeling patient-related information can improve the characterization of disease-associated white matter alterations when a sufficiently consistent imaging pattern is shared across patients. A key conceptual difference between TractSpLearn is that the sign of the resulting score carries meaningful information. In TractLearn, the magnitude of the deviation is the primary quantity of interest, and abnormalities can therefore be characterized using the absolute Z-score. In TractSpLearn, however, a positive score indicates that the observed tract profile is better explained by the patient-related representation than expected from the healthy reference distribution. Conversely, a negative score does not represent an abnormality of equal but opposite biological meaning. Therefore, only positive TractSpLearn Z-scores were considered as patient-related abnormality evidence in the present analysis.

This pattern was most evident in RHI, where TractSpLearn revealed significant group differences across a broader range of diffusion measures. Among the diffusion-derived parameters, the abnormalities identified by TractSpLearn were particularly prominent in AK and MD. This finding is consistent with previous diffusion MRI studies showing that both kurtosis- and diffusivity-based measures can be sensitive to microstructural alterations following TBI. For example, longitudinal studies of sports-related concussion have reported significant changes in MD after injury, with some abnormalities persisting beyond clinical symptom resolution (Wu et al., 2020). Diffusion kurtosis imaging has also demonstrated sensitivity to TBI-related white-matter alterations and may provide complementary information beyond conventional DTI measures (Stokum et al., 2015). A comparable observation was reported by Chung et al., who found significantly increased AK across widespread white matter in athletes with sports-related concussion, while conventional diffusion measures showed less extensive group differences (Chung et al., 2022). Similarly, Muftuler et al. demonstrated longitudinal changes in kurtosis-based measures following sports-related concussion, supporting the sensitivity of DKI-derived parameters to subtle and evolving white-matter alterations after head injury (Muftuler et al., 2020).

Importantly, a similar tendency was also observed in our exploratory analysis, where AK showed more widespread and more frequently observed abnormality patterns across subjects and tracts than several other diffusion parameters. In the RHI cohort, MD abnormalities were observed in a relatively limited number of tracts and generally at low cross-subject frequencies. The affected regions included left of fronto-pontine (FPT-Left, 10% of subjects) and Isthmus (CC6, 10%), as well as right of Superior longitudinal fascicle I, right of striato-parietal , and left of Thalamo-occipital (SLF_I.right, ST_PAR.right, and T_OCC.left ,5%). In comparison, AK showed a broader and more pronounced spatial distribution of abnormalities. Several of the regions identified by MD were also present in the AK results, including FPT_–left, SLF_I.right, ST_PAR.right, T_OCC.left, and CC6, with higher cross-subject frequencies observed for several of these tracts. In addition, AK revealed abnormalities in a number of other regions, most

prominently left of thalamo-precentral (T_PREC_left, 45%) and FPT_right (45%), right of cingulum (CG_right, 35%) and FPT_left (35%), as well as ST_PAR_left(5%) and rostral body premotor (CC3, 30%).

The involvement of these pathways is broadly supported by previous diffusion MRI studies of TBI. Yeh et al. reported longitudinal abnormalities involving the fronto-pontine tract, anterior thalamic radiation, corpus callosum, and other long projection pathways following TBI (Yeh et al., 2022). Palacios et al. also demonstrated diffusion abnormalities involving the cingulum in TBI and showed that alterations in this pathway were related to memory impairment (Palacios et al., 2011). Consistent with this, Wilde et al. reported reduced FA and increased diffusivity in bilateral cingulum bundles after TBI (Wilde et al., 2010). Corpus callosum involvement has likewise been repeatedly observed following TBI, and longitudinal diffusion studies have identified abnormalities in the callosum together with the cingulum and other major association pathways (Bendlin et al., 2008). More recently, Goubran et al. reported diffusion and kurtosis abnormalities involving the corpus callosum, superior longitudinal fasciculus, thalamic radiation, and cingulum in concussed football players (Goubran et al., 2023). These previous findings support the biological plausibility of the broader AK pattern observed here. Rather than simply representing stronger abnormalities in the same locations identified by MD, the additional involvement of fronto-pontine, cingulate, callosal, and thalamocortical-related pathways may indicate a different spatial pattern of white-matter microstructural alterations captured by AK.

The different performance observed between the RHI and PPCS cohorts further illustrates how the benefit of patient-informed modeling may depend on the consistency of disease-related imaging patterns. Although both TractLearn and TractSpLearn generally showed higher mean abnormality evidence in patients than in healthy controls, the relative patterns across the two patient cohorts differed between the methods. With TractLearn, the mean abnormality levels observed in RHI and PPCS were relatively similar, suggesting that a deviation-based framework primarily captures the extent to which an individual differs from the healthy reference, with limited sensitivity to differences in the underlying patient-related pattern. In contrast, TractSpLearn showed a clearer distinction between the two patient cohorts, with substantially stronger patient-related abnormality evidence in RHI than in PPCS. Nevertheless, TractSpLearn identified abnormality evidence in more PPCS patients than TractLearn, although the mean number of affected tracts among abnormality-positive patients decreased. These findings suggest that explicitly incorporating patient-related information can improve the characterization of white-matter abnormalities in both cohorts, while highlighting cohort-specific abnormality patterns when a more consistent disease-related imaging signature is present.

The more pronounced benefit observed in RHI may be related to differences in the underlying patho-

logical and imaging heterogeneity of the two cohorts. Persistent post-concussive symptoms can arise from heterogeneous pathophysiological mechanisms, and both the location and magnitude of diffusion abnormalities may vary considerably across individuals (Bouix et al., 2013). Such variability can reduce the consistency of the patient-related pattern available for model learning, particularly in a relatively small cohort, and may explain why the improvement in PPCS was more modest and why substantial overlap with healthy controls remained. Previous diffusion MRI studies of repetitive head impacts have demonstrated white-matter microstructural alterations associated with cumulative head-impact exposure (Bazarian et al., 2014; Brett et al., 2021). One possible explanation for the greater benefit observed in RHI is that more spatially extensive abnormalities increase the likelihood that some white matter regions are affected across multiple patients. These shared regional features could facilitate learning of the patient-related component in TractSpLearn, even when the overall abnormality patterns differ between individuals. However, this explanation remains hypothetical, as differences in the spatial overlap of abnormalities between RHI and PPCS were not directly assessed in the present study.

The individual spatial maps provide further support this interpretation by demonstrating substantial inter-subject variability in both healthy controls and patients. Even in RHI, where the group-level separation was clearer, the spatial distribution and extent of abnormalities differed considerably between individuals. Some healthy controls showed supra-threshold values, whereas some patients exhibited relatively limited abnormality evidence. These findings should not necessarily be interpreted as false-positive or false-negative lesion detection, because the Z-score maps represent model-derived diffusion abnormalities rather than histologically or clinically confirmed lesions. Their magnitude and spatial distribution may also be influenced by individual anatomical variability, diffusion measurement noise, tract reconstruction, spatial normalization, and variability within the healthy reference population. Accordingly, the individual maps should primarily be interpreted as visualizations of subject-specific abnormality evidence, whereas the systematic advantage of TractSpLearn is more appropriately assessed from the quantitative group-level analyses.

Taken together, the group-level and individual-level findings indicate that TractSpLearn does not simply increase abnormality scores across all patients. Rather, its advantage appears to depend on the degree to which a reproducible patient-related imaging pattern is present. This interpretation is consistent with the stronger group-level differentiation and broader AK abnormality pattern observed in RHI, while the PPCS findings, including abnormalities detected in more patients but fewer affected tracts per abnormality-positive patient, suggest that the framework may also provide additional value in more heterogeneous patient populations, although to a lesser extent.

Several limitations should be considered though. First, the patient cohorts are relatively small, particularly

for PPCS, which may limit the diversity and stability of the patient-related patterns learned by the model. Second, no independent lesion-level ground truth was available. Therefore, supra-threshold Z-score regions should be interpreted as model-derived abnormality evidence rather than confirmed pathological lesions, and the present study does not provide a formal assessment of diagnostic sensitivity or specificity. Third, the framework depends on several upstream processing steps, including diffusion model fitting, tract reconstruction, and spatial normalization. Although the same processing pipeline was applied to all subjects, methodological choices at these stages may introduce systematic, pipeline-dependent biases and thereby influence the sensitivity and spatial distribution of the detected abnormalities. Finally, each diffusion parameter was modeled separately in the present study, therefore, relationships between different parameters were not explicitly modeled. Future work could investigate whether joint modeling of multiple diffusion parameters, with a single multivariate model could better capture subtle and heterogeneous white-matter abnormalities.

5 Conclusion

Overall, TractSpLearn extends conventional tract-based normative modeling by incorporating patient-related information into individualized abnormality estimation. Rather than describing an unseen subject solely according to the magnitude of deviation from healthy controls, the proposed framework evaluates whether the observed tract profile exhibits increased preference for a patient-related representation relative to the healthy reference distribution. TractSpLearn improved the characterization of patient-related abnormalities in both RHI and PPCS cohorts, with the most pronounced group-level differentiation observed in RHI. In PPCS, the framework also identified abnormality evidence in more patients than TractLearn. These findings support the potential value of patient-informed modeling for individualized characterization of subtle and spatially distributed white-matter abnormalities, while also indicating that its benefit may depend on the consistency and heterogeneity of disease-related imaging patterns within the patient population.

Author Contributions

Jiqing Huang– Conceptualization; Methodology; Software (algorithm innovation and implementation); Formal analysis; Visualization; Writing – original draft.

Yi Chen– Methodology (algorithm design); Writing – review & editing.

Laurent Lamalle– Writing – review & editing.

Ali Al-Husseini– Investigation (data acquisition for the repetitive head injuries cohort); Data curation; Resources.

Anna Gard– Investigation (data acquisition for healthy controls and individuals with persistent post-concussive syndrome); Data curation; Resources.

Niklas Marklund– Resources (data provision); Methodology (protocol design); Writing – review & editing.

Markus Nilsson– Validation; Conceptualization (initial); Writing – review & editing.

Mohamed Ali Bahri –Validation; Statistical analysis; Writing – review & editing.

Christophe Phillips– Conceptualization; Methodology; Project administration (algorithm and manuscript); Supervision; Writing – review & editing.

Evgenios N. Kornaropoulos– Conceptualization; Methodology; Supervision; Writing – review & editing.

Acknowledgements

This work was supported by the FRS-FNRS and computational resources from the CÉCI (Consortium des Équipements de Calcul Intensif, <https://www.ceci-hpc.be/> platform and the SEGI (Service Général de l'Informatique) at the University of Liège (ULiège). We are especially grateful to Prof. David Colignon for his assistance with resource allocation and technical support.

Supplementary Material

The code for TractSpLearn is available on: <https://github.com/CyclotronResearchCentre/TractIGC>

References

- Attyé, A., Renard, F., Baciú, M., Roger, E., Lamalle, L., Dehail, P., Cassoude-salle, H., & Calamante, F. (2021). Tractlearn: A geodesic learning framework for quantitative analysis of brain bundles. *NeuroImage*, 233, 117927.
- Bazarian, J. J., Zhu, T., Zhong, J., Janigro, D., Rozen, E., Roberts, A., Javien, H., Merchant-Borna, K., Abar, B., & Blackman, E. G. (2014). Persistent, long-term cerebral white matter changes after sports-related repetitive head impacts. *PloS one*, 9(4), e94734.

- Bendlin, B. B., Ries, M. L., Lazar, M., Alexander, A. L., Dempsey, R. J., Rowley, H. A., Sherman, J. E., & Johnson, S. C. (2008). Longitudinal changes in patients with traumatic brain injury assessed with diffusion-tensor and volumetric imaging. *Neuroimage*, 42(2), 503–514.
- Bouix, S., Pasternak, O., Rathi, Y., Pelavin, P. E., Zafonte, R., & Shenton, M. E. (2013). Increased gray matter diffusion anisotropy in patients with persistent post-concussive symptoms following mild traumatic brain injury. *PloS one*, 8(6), e66205.
- Brett, B. L., Koch, K. M., Muftuler, L. T., Budde, M., McCrea, M. A., & Meier, T. B. (2021). Association of head impact exposure with white matter macrostructure and microstructure metrics. *Journal of Neurotrauma*, 38(4), 474–484.
- Choi, S., Cunningham, D. T., Aguila, F., Corrigan, J. D., Bogner, J., Mysiw, W. J., Knopp, M. V., & Schmalbrock, P. (2011). Dti at 7 and 3 t: Systematic comparison of snr and its influence on quantitative metrics. *Magnetic resonance imaging*, 29(6), 739–751.
- Chung, S., Chen, J., Li, T., Wang, Y., & Lui, Y. (2022). Investigating brain white matter in football players with and without concussion using a biophysical model from multishell diffusion mri. *American Journal of Neuroradiology*, 43(6), 823–828.
- Clemente, A., Attyé, A., Renard, F., Calamante, F., Burmester, A., Imms, P., Deutscher, E., Akhlaghi, H., Beech, P., Wilson, P. H., et al. (2023). Individualised profiling of white matter organisation in moderate-to-severe traumatic brain injury patients. *Brain Research*, 1806, 148289.
- Gard, A., Kornaropoulos, E. N., Portonova Wernersson, M., Rorsman, I., Blennow, K., Zetterberg, H., Tegner, Y., De Maio, A., Markenroth Bloch, K., Björkman-Burtscher, I., et al. (2024). Widespread white matter abnormalities in concussed athletes detected by 7t diffusion magnetic resonance imaging. *Journal of Neurotrauma*, 41(13-14), 1533–1549.
- Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., Nimmo-Smith, I., & Contributors, D. (2014). Dipy, a library for the analysis of diffusion mri data. *Frontiers in neuroinformatics*, 8, 8.
- Genç, B., Aslan, K., & İncesu, L. (2025). Early and persistent white matter damage in severe traumatic brain injury: A longitudinal diffusion tensor imaging analysis. *Current Research in MRI*, 3(3).
- Goubran, M., Mills, B. D., Georgiadis, M., Karimpoor, M., Mouchawar, N., Sami, S., Dennis, E. L., Akers, C., Mitchell, L., Boldt, B., et al. (2023). Microstructural alterations in tract development in college football and volleyball players: A longitudinal diffusion mri study. *Neurology*, 101(9), e953–e965.
- Hashim, E., Caverzasi, E., Papinutto, N., Lewis, C. E., Jing, R., Charles, O., Zhang, S., Lin, A., Graham, S. J., Schweizer, T. A., et al. (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, 97.

- Huang, S., Huang, C., Li, M., Zhang, H., & Liu, J. (2022). White matter abnormalities and cognitive deficit after mild traumatic brain injury: Comparing dti, dki, and noddi. *Frontiers in Neurology*, 13, 803066.
- Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magnetic resonance in medicine*, 76(5), 1574–1581.
- Kornaropoulos, E. N., Winzeck, S., Rumetshofer, T., Wikstrom, A., Knutsson, L., Correia, M. M., Sundgren, P. C., & Nilsson, M. (2022). Sensitivity of diffusion mri to white matter pathology: Influence of diffusion protocol, magnetic field strength, and processing pipeline in systemic lupus erythematosus. *Frontiers in neurology*, 13, 837385.
- Lee, A. L. (2020). Advanced imaging of traumatic brain injury. *Korean Journal of Neurotrauma*, 16(1), 3.
- Mahan, M. Y., Rafter, D. J., Truwit, C. L., Oswood, M., & Samadani, U. (2021). Evaluation of diffusion measurements reveals radial diffusivity indicative of microstructural damage following acute, mild traumatic brain injury. *Magnetic resonance imaging*, 77, 137–147.
- McInnes, L., Healy, J., & Melville, J. (2018). Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*.
- Muftuler, L. T., Meier, T. B., Keith, M., Budde, M. D., Huber, D. L., & McCrea, M. A. (2020). Serial diffusion kurtosis magnetic resonance imaging study during acute, subacute, and recovery periods after sport-related concussion. *Journal of neurotrauma*, 37(19), 2081–2092.
- Palacios, E. M., Fernandez-Espejo, D., Junque, C., Sanchez-Carrion, R., Roig, T., Tormos, J. M., Bargallo, N., & Vendrell, P. (2011). Diffusion tensor imaging differences relate to memory deficits in diffuse traumatic brain injury. *BMC neurology*, 11(1), 24.
- Roger, E., Attyé, A., Renard, F., & Baciú, M. (2022). Leveraging manifold learning techniques to explore white matter anomalies: An application of the tractlearn pipeline in epilepsy. *NeuroImage: Clinical*, 36, 103209.
- Stokum, J. A., Sours, C., Zhuo, J., Kane, R., Shanmuganathan, K., & Gullapalli, R. P. (2015). A longitudinal evaluation of diffusion kurtosis imaging in patients with mild traumatic brain injury. *Brain injury*, 29(1), 47–57.
- Veraart, J., Novikov, D. S., Christiaens, D., Ades-Aron, B., Sijbers, J., & Fieremans, E. (2016). Denoising of diffusion mri using random matrix theory. *Neuroimage*, 142, 394–406.
- Wasserthal, J., Neher, P., & Maier-Hein, K. H. (2018). Tractseg - fast and accurate white matter tract segmentation. *NeuroImage*, 183, 239–253. <https://doi.org/10.1016/j.neuroimage.2018.07.070>
- Wilde, E. A., Ramos, M. A., Yallampalli, R., Bigler, E. D., McCauley, S. R., Chu, Z., Wu, T. C., Hanten, G., Scheibel, R. S., Li, X., et al. (2010). Diffusion tensor imaging of the cingulum bundle in children after traumatic brain injury. *Developmental neuropsychology*, 35(3), 333–351.

- Wu, Y.-C., Harezlak, J., Elsaid, N. M., Lin, Z., Wen, Q., Mustafi, S. M., Riggen, L. D., Koch, K. M., Nencka, A. S., Meier, T. B., et al. (2020). Longitudinal white-matter abnormalities in sports-related concussion: A diffusion mri study. *Neurology*, *95*(7), e781–e792.
- Yeh, P.-H., Lippa, S. M., Brickell, T. A., Ollinger, J., French, L. M., & Lange, R. T. (2022). Longitudinal changes of white matter microstructure following traumatic brain injury in us military service members. *Brain communications*, *4*(3), fcac132.