A longitudinal diffusion tensor imaging study assessing white matter fiber tracts after sports related concussion
M. Murugavel¹, V. Cubon⁴, M. Putukian², R. Echemendia⁵, D. Osherson³, A. Dettwiler¹
¹Princeton Neuroscience Institute, ²University Health Services, ³Department of Psychology, Princeton University, NJ; ⁴Kent State University, Kent, OH; ⁵Psychological and Neurobehavioral Associates Inc., State College, PA

Annegret Dettwiler, Ed.D (corresponding author)
Associate Research Scholar
Princeton Neuroscience Institute
Princeton University
Washington Road
Princeton NJ 08544

Adjunct Assistant Professor
UMDNJ Robert Wood Johnson Medical School
New-Brunswick NJ 08903

phone: 609-258-9157
fax: 609-258-2574
email: adettwil@princeton.edu
Running title A longitudinal DTI study of WM tracts after SRC

Abstract
The extent of structural injury in sports related concussion is central to the course of recovery, long-term effects, and the decision to return to play. In the present longitudinal study, we used diffusion tensor imaging to assess white matter fiber tract integrity within two days, two weeks and two months of concussive injury. Participants were right-handed male varsity contact-sport athletes (20.2 ± 1.0 years of age), with a medically diagnosed sports related concussion (no loss of consciousness). They were compared to right-handed male varsity non-contact-sport athletes serving as controls (19.9 ± 1.7 years). We found significantly increased radial diffusivity in concussed athletes (n = 12, paired t-test, tract based spatial statistics, p < 0.025) at two days when compared to the two week post injury time point. The increase was found in a cluster of right hemisphere voxels, spanning the posterior limb of the internal capsule, the retrolenticular part of the internal capsule; the inferior longitudinal fasciculus, the inferior fronto-occipital fasciculus (sagittal stratum) and the anterior thalamic radiation. Post hoc, univariate, between group (controls vs. concussed) mixed effects analysis of the cluster showed significantly higher radial diffusivity at two days (p = 0.002) as compared to the controls; with a trend in the same direction at two months (p = 0.11). Results for fractional anisotropy in the same cluster showed a similar, but inverted pattern; fractional anisotropy was decreased at two days and at two months post injury when compared to normal controls. At two weeks post injury no statistical differences between concussed and control athletes were found with regard to either radial diffusivity or fractional anisotropy. These results support the hypothesis of increased radial diffusivity and reduced fractional anisotropy within 72 hours post injury, followed by recovery that may extend beyond 2 weeks. Radial diffusivity appears to be a sensitive measure of concussive injury.

Keywords Sports related concussion, mTBI, diffusion tensor imaging, radial diffusivity, longitudinal study
Introduction

The diagnosis of mild traumatic brain injury (mTBI) is often hindered by exclusive reliance on neurocognitive and clinical symptoms based on patient self-report. A more promising approach is to exploit radiological evidence from MRI and CT. Conventional clinical imaging techniques used to exclude intracranial hemorrhage or skull fracture, do not have the sensitivity to identify alterations in the neural microstructure resulting from mTBI. Advanced neuroimaging techniques, in particular, diffusion tensor imaging (DTI) are therefore worth exploring. The present study reports the use of DTI to assess White Matter (WM) fiber tract integrity in the brains of college athletes who sustained a Sports Related Concussion (SRC), one source of mTBI.

Sports are indeed a major cause of mTBI (often called “concussions”). A study by the Centers for Disease Control and Prevention estimates that 300,000 SRC’s occur annually in the United States (Thurman et al., 1998). However, this study only included concussions for which the person reported loss of consciousness, which is thought to characterize only a fraction of SRCs (Collins et al., 2003; Schulz et al., 2004). Given that athletes often do not report their injury, a more accurate approximation may be that 1.6 to 3.8 million SRCs occur each year including concussions for which no medical treatment is sought (Langlois et al., 2004).

According to the most recent consensus (McCrory et al., 2013), typical concussive injury results in the rapid onset of short-lived impairment of neurological function that resolves spontaneously. The authors of this statement affirm that: “a concussion may result in neuropathologic injury, but the acute clinical symptoms largely reflect a functional disturbance rather than structural injury.” This claim is questionable in light of recent neuroimaging research. Although clinical and cognitive symptoms may subside after approximately two weeks in most concussed athletes, neurological alterations can persist. For example, magnetic resonance spectroscopy (MRS) studies have demonstrated neurometabolic changes after SRC lasting up to 1 month post-injury (Chamard et al., 2013; Henry et al., 2010; Vagnozzi et al., 2010, 2013). Similarly, in a functional magnetic resonance imaging study, hyperactivation of the dorsolateral prefrontal cortex was found to persist beyond two months post injury in athletes whose symptoms subsided at two weeks after injury (Dettwiler et al., 2014). DTI studies demonstrating structural changes from repetitive concussive head impacts have been reported in ice hockey players over the course of a single season (Koerte et al., 2012b), in athletes with prolonged symptoms (Cubon et al., 2011) as well as in adolescents exhibiting close-to-normal scores on the Sports Concussion Assessment Tool (SCAT2) (McCrory et al., 2009) at two months post injury (Virji-Babul et al., 2013).

Compared to standard MRI, DTI offers a more sensitive assessment of focal ischemic lesions and diffuse axonal damage (Horsfield et al., 1998). Specifically,
DTI provides information about the WM microstructure and fiber tract integrity by measuring the Brownian motion of water molecules in the brain (Basser and Jones, 2002; Le Bihan et al., 2001; Johansen-Berg and Rushworth, 2009). Diffusion properties of water in tissue can be either isotropic or anisotropic. In tissues with isotropic diffusion, water molecules diffuse equally in all directions. Isotropic diffusion is typically found in the gray matter of the brain. In the anisotropic case water has a preferred direction of diffusion. Anisotropic diffusion is typically found in tissue with strong directional organization such as the deep WM where axons form tightly packed fiber bundles. In such tissue, diffusion is normally highly restricted along the fiber membranes. Measures of anisotropy thus provide information about the WM microstructure and WM fiber tract integrity (Basser and Jones, 2002; Le Bihan et al., 2001; Johansen-Berg and Rushworth, 2009), which is undetectable by conventional MRI methods.

DTI allows information about multiple diffusion gradients in a given tissue to be combined. A derived measure known as “fractional anisotropy” (FA) can then be used to quantify the degree of preferred diffusion direction in each voxel (Pierpaoli et al., 1996). Overall diffusion in a tissue is measured by “mean diffusivity” (MD), which is calculated as the mean of the three eigenvalues of the diffusion tensor (Mori, 2007). The eigenvalues of each directional vector can also be examined independently. The eigenvalue of the first eigenvector (also referred to as parallel diffusivity) was selectively altered in the presence of acute axonal damage in retinal ischemia in mice (Song et al., 2003). Similarly, “radial diffusivity” (RD), the mean of the second and third eigenvalues (Johansen-Berg and Rushworth, 2009) may be selectively sensitive to alterations of the myelin sheath, as demonstrated in an animal model (Song et al., 2002) and more recently in optic neuritis in humans (Naismith et al., 2009). These findings lend support to the sensitivity of diffusion measures with regard to specific pathologies.

Decreased FA has been reported in mTBI patients with a Glasgow coma score (GCS) of 13-15 within 24 hours post injury (Arfanakis et al., 2002). WM abnormalities have also been shown (Lipton et al., 2008) in patients with mTBI exhibiting persistent cognitive impairment (eight months to three years post injury). The latter investigators demonstrated decreased FA and increased MD in the corpus callosum, bilateral capsula interna and other subcortical WM structures. Significant correlations between decreased FA (corpus callosum, capsula interna, and centrum semiovale) within 10 days post injury and neuropsychological (NP) test scores obtained at 6 months post injury have been reported as well (Miles et al., 2008). Abnormalities of WM microstructure in mTBI patients with persistent cognitive impairment have been found (Niogi et al., 2008a) in the anterior corona radiata, the uncinate fasciculus, corpus callosum, inferior longitudinal fasciculus, and the cingulum bundle. Furthermore, significant correlations between attentional control and FA were found within the left anterior corona radiata as well as memory performance and FA within the uncinate fasciculus (Niogi et al., 2008b). A DTI study on patients with mTBI (Mayer et al.,
2010) demonstrated increased FA and decreased RD in the subacute phase after injury and subsequent partial normalization of FA values in left corona radiata and splenium. These studies provide evidence that anisotropy measurements cannot only be used to assess alterations in the microstructure of the WM but also provide a biomarker of cognitive function and dysfunction. Such markers may prove critical in refining the diagnosis, prognosis and management of mTBI. It should, however, be emphasized that the DTI studies using measurements of anisotropy discussed so far include diverse individuals with mild TBI and a GCS ranging between 13-15; athletes were not targeted for investigation. Although the mechanism of injury in SRC is believed to be comparable to non sports related mTBI, SRC represents the mildest form of mTBI. Most individuals with SRC will not score below 15 on the GCS, but will present with rapid onset of short-lived neurological impairment; they typically show no structural changes in traditional MRI and CT scans. It therefore seems prudent to exploit DTI technology to separately examine the case of SRC, especially given the prevalence of this condition (see above).

Only a few studies have assessed structural changes in adult athletes with SRC who do not score below 15 on the GCS. Increased RD and axial diffusivity after repetitive concussive head impacts in adult ice hockey players over the course of a single season were observed in the right precentral region, corona radiata and the anterior, posterior limb of the internal capsule (Koerte et al., 2012b). Decreased FA (in temporo-occipital WM) and lower cognitive function (CogState; Maruff et al., 2009) were found to be associated with high frequency heading rate (> 885–1800 headings per year) in amateur soccer players (Lipton et al., 2013). In college athletes exhibiting prolonged symptoms after SRC, increased MD has been reported in parts of the left inferior/superior longitudinal and fronto-occipital fasciculi, the retrolenticular part of the internal capsule, posterior thalamic and acoustic radiations (Cubon et al., 2011). Persistent microstructural alterations in deep WM have been shown in female contact sports athletes at 7 months post injury (Chamard et al., 2013); all participants were symptom free at this point of their recovery, suggesting that in female athletes, structural recovery may lag behind behaviorally assessed recovery by up to 7 months post injury. Finally, changes in WM microstructure were observed in a cohort of contact sports athletes with subconcussive blows to the head (26-399 hits), whereas no such changes were identified in six control participants (Bazarian et al., 2012).

There is thus growing evidence suggesting that even in the absence of clinically symptomatic concussions (Bazarian et al., 2012; Koerte et al., 2012a; Lipton et al., 2013) or at a stage of recovery when athletes are symptom free (Chamard et al., 2013) they are likely to exhibit WM alterations when advanced neuroimaging techniques are used to examine their brains. These findings suggest, that DTI may be a useful imaging tool to assess the severity of a concussion and may provide a biomarker for structural injury. DTI examination of the brain may thus serve to monitor the reorganization and reversal of WM injury, and to predict recovery. The aim of the present study was to track changes of WM fiber tract
integrity during the two months following SRC using advanced DTI.

**Materials and Methods**

**Participants**

All concussed participants were varsity level college students enrolled in the Princeton University concussion program for high risk sports. The program ensures systematic documentation of athletic history, physical exam, and baseline NP testing including SCAT2 (McCrorry et al., 2009) and Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) (Lovell et al., 2007). Princeton’s concussion program also provides acute care and long term monitoring. All athletes involved in this study were diagnosed with a concussion by team physicians using criteria outlined by the 4th International Consensus Conference on Concussion in Sport (McCrory et al., 2013). Post-injury testing included SCAT2, traditional “paper and pencil” NP tests, ImPACT, the Patient Health Questionnaire (PHQ-9) (Kroenke and Spitzer, 2002) and the Generalized Anxiety Disorder (GAD-7) Questionnaire (Spitzer et al., 2006). The PHQ-9 and the GAD-7 are assessments for depression and generalized anxiety respectively. The baseline and post-injury testing protocols were identical to those described in our earlier publication (Dettwiler et al., 2014).

Following their most recent concussion, a certified athletic trainer and team physician at the University Health Services evaluated athletes within 48 hours post injury. None of the athletes experienced a loss of consciousness and their overall symptomatology did not warrant further assessment by the Glasgow Coma Scale or the use of a clinical radiological exam. All concussed athletes underwent NP testing within 24-48 hours after injury. Abnormal NP performance was determined through comparison of post injury NP scores to the athlete’s baseline scores. Specifically, abnormality of ImPACT clinical composites was based on Reliable Change Indices at the 0.8 Confidence Interval (Iverson et al., 2003; Lovell et al., 2007). Similarly, scores on the traditional NP test performance were examined using Princeton specific normative data. Data from both ImPACT and the NP test were integrated and interpreted by an experienced clinical neuropsychologist.

Athletes were kept out of activity until they were symptom free and their clinical exam including balance and NP evaluations were considered to have returned to baseline levels. Return to activity decisions were made by the team physician, who supervised a personalized return to play progression that exposed athletes to gradual increases in physical exertion as per the 1st International Consensus Conference on Concussion guidelines (Aubry et al., 2002). Athletes were cleared to return to full contact play once they were symptom free at rest, had successfully completed the exertional program and were neurocognitively functioning at baseline levels.
A total of 21 right-handed, male, varsity level contact sport athletes (mean age 20.19, S.D. 1.03, age range 18-22 years) who suffered a SRC were enrolled in the study. In addition to having no contraindications to MR imaging, participants had no self reported history of medical, genetic, or psychiatric disorder. History of concussion was obtained through self-report. It should be noted that under-reporting of concussion by athletes has been suggested in previous studies (Torres et al., 2013). An objective evaluation of the number of previous concussions in contact sport athletes is therefore difficult. Among the pool of 21 concussed athletes, 12 reported no prior history of concussion, five reported one prior concussion; three reported two prior concussions and one reported three prior concussions. The mean time since the last self reported concussion for the latter nine concussed athletes was 2.75 years, SD 3.02 years; see Table 1.

Table 1: Demographics of controls and concussed athletes

<table>
<thead>
<tr>
<th>Controls/Concussed</th>
<th>MRI Inclusion</th>
<th>Age (years)</th>
<th>Sport</th>
<th># Prior concussions</th>
<th>NP Normal post injury (days)</th>
<th>Symptom free (days)</th>
<th>Return to play (days)</th>
<th>2 days scan</th>
<th>2 wks. scan</th>
<th>2 mon. scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Age (years)</td>
<td>Sport</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>Crew</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>Squash</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>Crew</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>Crew</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>Track + Cross country</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>18</td>
<td>Crew</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>21</td>
<td>Track</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>21</td>
<td>Volleyball</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>22</td>
<td>Track + Cross country</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>22</td>
<td>Cross country</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>19</td>
<td>Crew</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>Track</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>19</td>
<td>Track</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>18</td>
<td>Swimming</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>19</td>
<td>Volleyball</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>22</td>
<td>Cross country</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>21</td>
<td>Ice hockey</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>20</td>
<td>Football</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>Basketball</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>19</td>
<td>Football</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>Sprint Football</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Returned to play after 1st injury at 31 days; sustained a 2nd concussion and decided not to return to play
* NP testing never reached normal range before athlete graduated/season ended
* Not normal at 2 weeks, not repeated until 2 months since season over
* Subjected discontinued from study and decided not to return to play (although cleared to do so for next, season)
* No return to play; graduated
Y: Scan included in analyses, M: Scan excluded due to motion exceeding threshold, X: Period of study during which scanner amplifier failed, D: Subject discontinued

Healthy control participants included 16 age matched, right-handed, male varsity non-contact athletes (mean age 19.9, S.D. 1.67, age range 18-22 years), with no contraindications to MR imaging and no self reported history of prior head trauma, psychiatric, neurological or developmental disorders. All athletes (concussed and controls) gave written consent to participate in the study, which was approved by the Princeton University’s Institutional Review Panel for Human Subjects Research. The concussed athletes were scanned at ~2 days, ~2 weeks and ~2 months post injury. The controls were scanned once. All athletes repeated SCAT2, PHQ-9, GAD-7 and NP testing assessments synchronized with the three imaging sessions of the concussed athletes. Concussed athletes participated in additional NP testing in between imaging sessions, as clinically indicated and requested by the team physician. There were 8 instances (during a single contiguous time period identified by ‘X’ in Table 1) when data collection was not possible on concussed athletes in the time interval required by the
experimental design of the present study due to hardware maintenance issues. There was also one instance of a concussed athlete deciding to discontinue participation in the study (identified by ‘D’ in Table 1). A strict data quality assurance protocol (described in the data pre-processing section) resulted in the exclusion of scans for 10 concussed and 2 controls (identified by ‘M’ in Table 1). In total, 14 controls, 16 concussed athletes at ~2 days, 17 concussed athletes at 2 weeks and 13 concussed athletes at 2 months were included in the analyses (identified by ‘Y’ in Table 1). From this pool of concussed athletes (see Table 1) only 12 concussed were imaged at both the 2 day time point and at 2 weeks, 11 concussed athletes were imaged at both 2 weeks and 2 months post injury.

**Imaging protocol**

Diffusion weighted images (single-shot spin echo pulse sequence with parameters adapted from Cubon et al., 2011) were acquired with a 16 channel, phase array coil (Siemens, Erlangen, Germany) on a whole body 3T Siemens ‘Skyra’ scanner; TR = 12100 ms, TE = 96 ms, 70 axial slices, voxel size $1.88 \times 1.88 \ mm^2$ in plane, slice thickness = 1.9 mm, field of view = 256 mm, 64 gradient directions, b-value 1000 s/mm$^2$, 8 volumes with no diffusion weighting (b = 0), 2 runs, yielding a total scan time of 26 minutes 52 seconds. In order to facilitate image volume registration to the Montreal Neurological Institute (MNI) space, a high resolution T1 weighted MPRAGE image was additionally acquired at the start of each imaging session; TR = 1900 ms, TE = 2.13 ms, 192 sagittal slices, $0.90 \times 0.94 \times 0.94 \ mm^3$ voxel resolution, flip angle = 9 degrees, field of view = 240 mm and a total anatomical scan time of 4 minutes 26 seconds. Care was taken to minimize subject motion with pre-scan instructions and comfortable neck padding. Participants watched a pre-selected, movie of their choice from an online streaming service during the entire scanning session.

**Data preprocessing and quality assurance**

All data processing was done within the FSL suite (version 4.1.9) (Smith et al., 2004). The two averages of the acquired diffusion weighted images of each subject were concatenated in the order of image acquisition and visually inspected for signal drop offs and other imaging artifacts. All acquired data passed visual inspection. Eddy current correction was done for each subject’s concatenated dataset, employing the first B0 volume for reference. Each volume’s registration parameters from the eddy correction step was then used to implement a strict, quantitative, quality assurance protocol based on recent findings (Ling et al., 2012a). Mean motion estimates (translation, rotation in three dimensions) were calculated for each group separately. All individual subject-scans with motion estimates greater than three standard deviations from the mean or scans with a net translational motion estimate exceeding two voxels were excluded in their entirety (10 scans of concussed athletes and 2 controls). The concatenated B vectors corresponding to the applied diffusion gradients were then corrected for motion (rotation component; see Leemans and Jones,
2009) before FSL function ‘dtifit’ (Smith et al., 2006) was applied to fit a diffusion tensor model (Basser et al., 1994), generating the three principal eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$ at each voxel. This step additionally provides scalar diffusion measures of WM microstructure, $\text{FA} = \sqrt{((\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2)/(2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2))}$ and $\text{AD} = \lambda_1$. $\text{RD} = [\lambda_2 + \lambda_3]/2$ and $\text{MD} = [\lambda_1 + \lambda_2 + \lambda_3]/3$ volumes were generated using the radial eigenvalues $\lambda_2$, $\lambda_3$.

### Statistical analyses

Between group T-tests, using the function ‘randomise’, were performed via Tract Based Spatial Statistics (TBSS, FSL version 4.1.9, Smith et al., 2006) on the skeletonized WM fiber tracts for all derived scalar diffusion measures of WM microstructure FA, AD, RD and MD. All TBSS processing steps followed recommended guidelines (Smith et al., 2006). The ‘FMRIB58_FA_1mm’ image volume in MNI space, included in FSL version 4.1.9, served as the target for initial non-linear registration (Andersson et al., 2007) of subject FA volumes. The mean WM skeleton based on the included participants FA volumes was thresholded to only include voxels with $\text{FA} > 0.25$ in order to restrict the analyses to the core WM tracts. The brain stem and the cerebellum were removed (mask included in the FSL suite 4.1.9) from all analyses since individual subject variability in brain volumes resulted in omission of inferior parts of these structures in a few cases. The number of ‘randomise’ permutations was set at 10000 with Threshold Free Cluster Enhancement (TFCE) option enabled (Smith and Nichols, 2009). Between sessions comparisons of the concussed athletes were made via paired T-tests (TBSS, TFCE, all permutations, variance smoothing of 2 voxel sizes) for all four diffusion measures of WM microstructure (FA, AD, RD and MD).

Two Post hoc tests, a traditional univariate Mixed Effects approach (Pinheiro and Bates, 2000) and a multivariate bootstrap method (Efron and Tibshirani, 1993) were selected to test if regions identified via a whole brain TBSS analysis differed (in terms of diffusion metrics of WM microstructure) between groups over time. The Mixed Effects model incorporates both ‘fixed effects’ and ‘random effects’ and is particularly useful in longitudinal studies because of its ability to deal with repeated measures and missing values (See Appendix A1 for details). The multivariate Bootstrap has the added advantage of accounting for combined responses of identified diffusion measures WM microstructure and is preferred in situations of moderate sample sizes such as this study. (See Appendix A2 for the algorithm). All post-hoc tests were run using the open source statistical software R (http://www.R-project.org).

### Results

No between group differences were found in TBSS analyses via pooled T-tests
at p < 0.05 (two sided), Family Wise Error (FWE) corrected with TFCE option enabled. Significant differences (pointing to structural alterations) were observed in the paired, between concussed sessions T-test (2 days vs. 2 weeks, p < 0.025, FWE corrected, TFCE) of the RD measure, with the cluster indicating greater RD values at 2 days as compared to 2 weeks. The significant RD cluster consisted of 469 contiguous voxels in standard space (MNI, FMRIB58_FA_1mm). The regions implicated are all in the right hemisphere; posterior limb of the internal capsule (IC), retrolenticular part of the IC, sagittal stratum (inferior longitudinal fasciculus & inferior fronto-occipital fasciculus) and anterior thalamic radiation. The John Hopkins University (JHU) ICBM-DTI-81 WM and JHU WM tractography atlases (Hua et al., 2008; Mori et al., 2008; Wakana et al., 2007) included in FSL version 4.1.9 were used to determine the anatomical regions referenced. Interestingly, these regions are almost identical to those reported earlier (Cubon et al. 2011) in the contralateral hemisphere. In addition, a trend (p<0.05, one sided, FWE, TFCE, two clusters for a total of 348 voxels in MNI space; FMRIB58_FA_1mm) was observed in the FA measure with both clusters overlaying about 58 % of the aforementioned significant RD cluster but with the result trending in the opposite direction. FA values were greater at 2 weeks as compared to the values at 2 days post injury. No paired, significant differences in FA, RD and AD measures were observed between sessions two and three or session one and three.

Figure 1: Results of the paired, between concussed session (two days vs. two weeks, corrected p < 0.025), TBSS T-test of the RD values on the WM skeleton. The voxels (inflated into adjoining local tracts for visualization) showing significantly higher RD values at two days as compared to two weeks have been highlighted by color mapping (red-yellow). These voxels have been overlaid onto their corresponding WM skeleton (green). The underlay is the ‘FMRIB58_FA_1mm’ image volume (grayscale).
Figure 2: Individual trajectories of the mean RD values downloaded from the paired TBSS T-test RD mask (two days vs. two weeks, p < 0.025 corrected) for all concussed athletes across all three sessions. The red circles indicate those concussed athletes who participated at all three sessions. Solid red lines link these circles. The red stars indicate those athletes who had imaging data at a maximum of two time points and connected by dashed red lines. The blue line marks the mean value of the controls while the solid black line connects the mean of the concussed across the three time points.

The RD voxel (paired TBSS, two days vs. two weeks, p<0.025, FWE corrected, TFCE) mask was used to download individual mean RD and FA values from all eligible subject's volumes in order to conduct post-hoc between group statistical tests. Fig. 2 illustrates the individual trajectories of the downloaded mean RD values.

The results of the between group, Mixed Effects analyses for RD and FA are presented in Table 2. The results of the Mixed Effects model suggest that RD values are on average significantly higher 2 days post injury (two sided p-value = 0.002) as compared to the controls, but the difference at 2 months represents more of a trend (two sided p-value = 0.11). At two weeks post injury there is no statistical difference between the groups with regard to the RD measure. The FA results show a similar but inverted pattern. FA values are on average lower for the injured athletes at all three time points as compared to the controls. At two days post injury, FA values are significantly lower in the concussed as compared to the controls (two sided p-value = 0.0008). At two months the differences persist (two sided p-value = 0.044), but at two weeks the average difference from controls is not statistically significant. The results of the multivariate (FA and RD) Bootstrap analysis are presented in Table 3. These results point to significant differences between groups at two days and a trend at two months.
### Table 2: Between group Mixed effects analysis of FA and RD

<table>
<thead>
<tr>
<th></th>
<th>FA Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>one-tailed p value</th>
<th>two-tailed p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.619967</td>
<td>0.004693</td>
<td>32.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 days</td>
<td>-0.030114</td>
<td>0.008278</td>
<td>-3.64</td>
<td>0.0004</td>
<td>0.0008*</td>
</tr>
<tr>
<td>2 wks.</td>
<td>-0.007985</td>
<td>0.007536</td>
<td>-1.06</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td>2 mon.</td>
<td>-0.017821</td>
<td>0.008550</td>
<td>-2.08</td>
<td>0.022</td>
<td>0.044*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RD Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>one-tailed p value</th>
<th>two-tailed p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>4.48E-04</td>
<td>4.56E-06</td>
<td>98.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 days</td>
<td>3.06E-05</td>
<td>9.21E-06</td>
<td>3.32</td>
<td>0.001</td>
<td>0.002*</td>
</tr>
<tr>
<td>2 wks.</td>
<td>3.90E-06</td>
<td>8.07E-06</td>
<td>0.48</td>
<td>0.32</td>
<td>0.64</td>
</tr>
<tr>
<td>2 mon.</td>
<td>1.53E-05</td>
<td>9.54E-06</td>
<td>1.6</td>
<td>0.059</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Table 2: Between group Mixed effects analysis of FA and RD

<table>
<thead>
<tr>
<th></th>
<th>Hotelling $T^2$ statistic for 3 time points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 days</td>
</tr>
<tr>
<td>$T^2$</td>
<td>8.253</td>
</tr>
<tr>
<td>Boot p value</td>
<td>0.0273*</td>
</tr>
<tr>
<td>Percentiles of Null Bootstrap distribution of $T^2$</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>6.463</td>
</tr>
<tr>
<td>97.5%</td>
<td>8.469</td>
</tr>
<tr>
<td>99%</td>
<td>10.988</td>
</tr>
</tbody>
</table>

Table 3: Results of the between group, multivariate analysis of FA, RD measures using the Bootstrap method for hypothesis testing.

Mean RD values from the significant RD cluster and its local vicinity within the WM tract i.e. the inflated RD cluster (see Fig. 1) were correlated with mean RD measures of the remaining deep WM clusters to assess if the same trend existed ‘globally’. The deep WM tracts for each individual volume in MNI space were masked by JHU ICBM-DTI-81 WM atlas. They were further constrained to include only WM by thresholding the corresponding FA volumes at 0.25. Two tailed p values (testing the null hypothesis of no correlation) of the control group was 0.14 and less than 0.05 for the concussed across all three imaging sessions. Fig. 3 illustrates the individual trajectories of the mean RD values from the aforementioned deep WM region.
Discussion

Results of the current study reveal structural alterations in the deep WM of the brain over the course of the two months following injury. Our primary finding is the significant difference observed between sessions one and two (two days vs. two weeks), within the concussed group in the paired TBSS t-test of the RD measure with greater values at two days as compared to two weeks. In addition, the same TBSS comparison revealed a reverse trend for the FA measure (within concussed session, paired TBSS t-test) with greater values at two weeks as compared to two days post injury with significant overlap of the FA with the RD cluster. The significant RD cluster, spans across parts of the posterior limb, the retro lenticular part of the IC, the inferior fronto-occipital fasciculus, inferior longitudinal fasciculus (sagittal stratum) and extends into the anterior thalamic radiation. Of specific note are two recent TBSS studies; a pilot study of veterans with combat related TBI (Kim and Jorge, 2010) and a comparable study in athletes with prolonged symptoms after SRC (Cubon et al., 2011), both of which reported nearly the same anatomic region in the contralateral hemisphere as compared to the significant RD cluster identified in this study. Other research, involving patients with a GCS 13-15, using a broad range of analyses, including
TBSS, have reported abnormal diffusion measures in a subset of regions covered by the significant RD cluster reported in the current study. Specifically, such regions were observed in the internal capsule (Arfanakis et al., 2002; Bazarian et al., 2007, 2012; Grossman et al., 2012; Huisman et al., 2004; Inglese et al., 2005; Lipton et al., 2012; Lo et al., 2009; Mayer et al., 2010; Miles et al., 2008), in either the inferior fronto-occipital and/or inferior long fasciculus (Geary et al., 2010; Messé et al., 2011; Nioi et al., 2008a,b; Smits et al., 2011) and the anterior thalamic radiation (Messé et al., 2011). The current study lends further support to the earlier hypothesis (Cubon et al. 2011), which suggested the prevalence of crossing and merging WM fiber tracts in the anatomic region of the RD cluster might make this particular area more vulnerable to the type of forces acting on the brain during the course of a concussion. This hypothesis posits that certain anatomic regions are more vulnerable to trauma than others, independent of the biomechanical load dynamics of the injury. Finite Element Method (FEM) based reconstructions (Viano et al., 2005) of head impacts from the National Football League found early strain ‘hot spots’ along the temporal lobe. These strain ‘hot spots’ then migrated to the fornix, midbrain and corpus callosum and were manifest in 9 out of 22 concussion reconstructions. A later FEM study (McAllister et al., 2011) correlating FA and MD values in a different ROI (corpus callosum) with computer simulations of the impact appear to show strain resulting in ‘hot spots’ in the temporal lobe, as well, although secondary in intensity to the corpus callosum. These findings provide additional support for the increased vulnerability of the anatomical regions of the significant RD cluster identified in the current study. Future studies might further elucidate the effect of impact forces by correlating injury mechanism and load dynamics to brain pathology (via post injury in-vivo imaging) with retrospective video analyses coupled to a head impact telemetry system (Guskiewicz et al., 2007).

Given the variability of patient characteristics and concussive injury mechanisms, one may question the validity of searching for common regions of pathology, which is inherent to any between group, voxel wise analyses of mTBI (Kim et al., 2013; Kou et al., 2010; Lipton et al., 2012). Instead, mTBI may have a unique spatial pattern of injury in each individual patient’s brain. Researchers taking this perspective compare the voxels of individual patients (diffusion measures) in standard space with the corresponding voxel set of a control group. Extreme voxels, deviating either positively or negatively from the control group, are then labeled and clustered (with multiple comparisons correction). The summary statistics of such abnormal loci reported in recent mTBI literature (Ling et al., 2012b; Lipton et al., 2012) reveal significant positive and negative clusters with significant between group differences (Kim et al. 2013). Future approaches to tracking recovery of individual concussions should compare the efficacy of the latter techniques against monitoring of diffusion measures over time, obtained from predetermined regions of vulnerability, such as the mask of the significant clusters arising in the current study.

No previous study has assessed the type of SRC examined here (with no LOC)
at three time points (2 days, 2 weeks and 2 months). Although our permutation tests on the whole brain WM skeleton did not reveal any significant between group differences, the comparisons of the voxels within the RD cluster showed significant between group difference at 2 days and a trend at 2 months. Closely related studies have demonstrated RD as a useful measure to assess the continuum of the mild end of TBI. RD values have been shown to increase (paired TBSS t-tests) over the course of a season in individual contact sport athletes (Koerte et al., 2012b) demonstrating significant increases only in Trace, AD and RD measures when comparing preseason with postseason images. The posterior limb of the IC was reported as a region (among others) with significant differences in structural measures between pre and post season, which incidentally is a region implicated in the current study. Furthermore a significant increase in RD was found in 3 athletes as compared to the rest of the players in the study, who sustained a medically diagnosed concussion during the course of the season. No significant difference was found in Trace, FA or AD. Another study (Koerte et al., 2012a) compared the WM integrity of swimmers to professional soccer players, with exposure to ‘headings’ (without a symptomatic concussion), and found increased RD in several areas including inferior fronto-occipital fasciculus (a region implicated in the current study). No significant differences were found in the FA and MD measures. These studies (Koerte et al., 2012a,b) suggest that RD might be a potentially sensitive measure to sub-concussive hits. A recent DTI study (Shin et al., 2013) on cerebral WM in 74 boxers and 81 mixed martial arts fighters found that a history of prior knockouts (the ‘knockout’ measure includes ‘technical knockouts’ with no subsequent LOC) could predict increased RD in the corpus callosum, isthmus of the cingulate gyrus, pericalcarine sulcus, the precuneus and the amygdala in the group of boxers. The same regions had increased MD and decreased FA values. The ‘knockout’ measure additionally predicted significantly increased RD in the posterior cingulate in the group of mixed martial arts fighters. In addition they found that the number of prior fights did not predict differences in diffusion measures, suggesting that diffusion measures were sensitive to potential sub-concussive hits or concussions, as opposed to time of exposure to the sport. In a longitudinal mTBI study (GCS 13-15) with imaging sessions at 24 hours, 1 week and 1 month post injury, statistical trends were reported (Zhu et al., 2010) in the paired between concussed session, based on TBSS t-tests of RD (greater at 24 hours vs. 1 month post injury) and FA (lower at 24 hours vs. 1 week post injury). It should, however, be noted, that the lack of significant differences might have been due to random assignment of participants to two different scanners. A region of interest study (Kumar et al., 2009) reported increased RD in a sample of mild and moderate TBI patients; imaging occurred an average of 8.9 days post injury. Despite the fact that these findings appear to lend support to the sensitivity of RD with regard to mTBI, future DTI studies should additionally assess the validity of RD as a diffusion measure for the assessment of mTBI.

The major finding of the current study is the occurrence of significant temporal changes in radial diffusivity between ~2 days and 2 weeks post injury in a sample
of concussed athletes. Multiple cross-sectional mTBI studies with one or more time points (Ling et al., 2012b; Singh et al., 2010; Zhu et al., 2010) have broadly discussed the coupled, inverse expression of RD/MD and FA measures in the acute and sub-acute phases post injury i.e. increased RD/MD and/or decreased FA or decreased RD/MD and/or increased FA with respect to matched controls. The results of the current study support an earlier hypothesis (Arfanakis et al., 2001) on the role of focal neurofilament misalignment, as an initializing mechanism leading to decreased FA, increased RD and reduced AD in human mTBI patients (GCS 13-15) imaged around 24 hours post injury (Arfanakis et al., 2001; Singh et al., 2010). Such misalignment had been observed to be manifest within 6 hours of axonal injury in animal models (Christman et al., 1994; Grady et al., 1993; Pettus et al., 1994; Povlishock and Christman, 1995). While the increased RD/MD and/or decreased FA mode is frequently reported in mTBI as well as moderate/severe TBI literature (see review of Shenton et al., 2012) and in studies of sub-concussive hits (Bazarian et al., 2012), there is a lack of consensus on the broad directionality of the diffusion measures after a concussive injury (Bazarian et al., 2007; Mayer et al., 2010; Wilde et al., 2008). Earlier findings of increased FA and reduced RD following mTBI (Mayer et al., 2010), have been replicated (Ling et al., 2012b). The authors reported a significant reduction in both the count and the volume of positive clusters representing regions of high FA over a 4-month period with the corresponding reduction in self-reported symptomatology suggesting recovery. Cytotoxic edema (Bazarian et al., 2007) was suggested as a potential explanation for the increased FA findings during the recovery interval. A recent longitudinal study (Lipton et al., 2012) assessed individual FA abnormalities in mTBI patients at ~2 weeks, 3 months and 6 months post injury (analyses elaborated in Kim et al., 2013 and similar to Ling et al. 2012b). They found that the count of low FA voxels decreases at both 3 and 6 months, but the count of high FA voxels increased at 3 months followed by a decrease at 6 months as compared to their initial assessments at 2 weeks post injury. The authors note that the continued expression of the positive clusters is inconsistent with cytotoxic edema, which drives ionic edema and signals a premorbid cellular process leading to necrotic cell death (Liang et al., 2007). In discussing these findings, other researchers (Shin et al., 2013) suggest the possibility that contact sport athletes represent a distinct population due to their continued exposure to sub-concussive hits leading to constant WM injury and recovery cycles and therefore might present a different recovery profile from the civilian, non-contact sport population suffering a single mTBI episode. It must be noted that at least one study on SRC (Henry et al., 2011) showed significantly higher FA, AD and lower MD (as compared to non contact controls) values at two time points; ~81 hours (on average) and 6 months post injury, suggesting no significant recovery in diffusion measures during that time interval. Future work is needed to address these observed differences of diffusion metrics during recovery after SRC.

Animal models of TBI additionally support the findings of the present study. For example, a recent controlled cortical impact study on rats (Budde et al., 2011) showed significantly increased RD and decreased FA in WM. RD may also be
selectively sensitive to alterations of the myelin sheath (Johansen-Berg and Behrens, 2009) as shown in the mouse model (Song et al., 2002, 2005) and more recently in optic neuritis (Naismith et al., 2009). Recovery as observed by histology after controlled cortical impact induced TBI in a rat model have correlated with increases in FA (Ding et al., 2013; Jiang et al., 2011); this has been attributed to axonal recovery and increased oligodendrocyte generation. A recent histology study scaling biomechanical loads to approximate mTBI in swine found axonal swellings and an accumulation of neurofilament protein (Browne et al., 2011). These observations could be expected to increase RD and lower FA according to the focal neurofilament misalignment hypothesis discussed earlier (Arfanakis et al., 2001). Further evidence is needed to confirm these findings in humans.

A traditional interpretation of FA increases from 2 days to 2 weeks post injury and corresponding decreases in RD would indicate that patients are recovering from mTBI. This interpretation has been proposed in more severe TBI (Sidaros et al., 2008). The fact that no differences (in all four diffusion metrics considered) were identified between two weeks and two months in the current study, might in part be due to inter subject variability of these measures. The finding of significant between group differences of the cluster at two days provides support for the view that diffusion measures may offer the required sensitivity to assess injuries as mild as the ones examined in this study. Diffusion measures at the identified anatomic location might have future diagnostic potential as a signature of concussion. Individual subject baselines or a database of normative values in the early phase of concussion might allow for identification of athletes at greater risk of prolonged recovery. There were no significant between group differences at two weeks. While this could be interpreted to be indicative of recovery, it should be noted that inter subject variability could potentially mask an ongoing or unresolved recovery process at two weeks. A future study should include a baseline MRI scan and a time point at one month to further clarify the course of the recovery process, exhibited through diffusion abnormalities.

A majority (80 - 90%) of concussions resolve between 7-10 days post injury as measured by behavioral assessments (McCrory et al., 2009). However the results of the present study provide evidence of neural recovery extending to at least 2 weeks from a structural perspective. Although these data do not inform us about the absolute maxima and minima of the diffusion metric trajectories due to the absence of measurements between two weeks and two months post injury, the statistical trend detected via the between group analyses at two months, suggests a minor relapse in the recovery of the structural measures of WM integrity. This finding, taken together with the observed variability in the trajectories of RD and FA between two weeks and two months, might be reflective of the athletes' exposure to sub-concussive hits following return to play (see earlier description of Koerte et al., 2012a,b; Shin et al., 2013). A more recent study (Marchi et al., 2013) and the first to relate diffusion measures to biomarkers in athletes with sub-concussive hits, reported a positive correlation between the percentage change in football post minus preseason levels of serum
auto-antibodies of the astrocytic protein S100B (considered a peripheral marker of blood brain barrier dysfunction) and the percentage of voxels with changes in MD during the corresponding time period. The same study reported a significant positive trend between the Head Hit Index (a derived measure accounting for both the number and severity of sub-concussive hits during a single game) and the increased post game (as compared to baseline) S100B levels of individual athletes. In addition, these, significant post game increases in S100B levels were detected only in athletes with sub-concussive hits (confirmed via game video analyses). These findings suggest that subject specific exposure to sub-concussive hits after return to play may be a potential factor affecting recovery of diffusion measures (Bazarian et al., 2012; Marchi et al., 2013).

Furthermore, variability in the trajectories of the diffusion measures might be affected by differences in number of prior concussions (Shin et al., 2013), timing of each athlete’s return to play (see Table 1) and individual genetic predisposition (Waters and Nicoll, 2005; Waters et al., 2013). Experimental designs of future studies should include the assessment of subconcussive hits, extending at least to the end of season.

Conclusions
This is the first longitudinal study that tracks diffusion measures of contact sport athletes following a single episode of SRC with no LOC at ~2 days, 2 weeks and 2 months. This study provides support for the hypothesis of increased RD and reduced FA within 72 hours post injury followed by patterns of recovery. It further suggests that neural recovery may extend beyond 2 weeks as described in other similar imaging studies (Dettwiler et al., 2014; Vagnozzi et al., 2010). RD was found to be sensitive marker of SRC with potential for personalized imaging based diagnosis.

Acknowledgments
We would like to acknowledge the athletic trainers of University Health Services for their assistance with subject recruitment and NP testing. This work was funded by the New Jersey Commission for Brain Injury Research grant No. 10-3217-BIR-E-0, AMSSM Foundation grant No. 005548, the Goldstein Family Fund, and the Peter & Cynthia Kellogg Foundation.

Author Disclosure Statement
No competing financial interests exist.

Appendix
A1. Mixed Effect Model
The Mixed Effect model was selected based on the following observations unique to this experimental design:
1. ‘Participant’ is a random effect and there are repeated measures over the same individuals. The repeated measures are unbalanced due to the missing values. This model also accounts for any correlations between the scalar diffusion measures of WM microstructure.

2. Concussed participants cannot be matched (paired) to controls as controls are imaged at only one time point.

3. This model treats time as a fixed effect, with time = 0, denoting the controls. It then allows for comparison of the concussion effect at time = 2 days, 2 weeks and 2 months in relation to the controls.

The model was implemented using the lme4 library in R, specifically, using the ‘lmer’ function.

A2. Multivariate analysis using Bootstrap method for hypothesis testing

The following procedure describes the bootstrap method employed for testing the mean effect at 2 days, 2 weeks and 2 months with respect to the controls. This test preserves repeated measures, the missing value structure and any correlation structure between the scalar diffusion measures of WM microstructure.

Step 1. Use the data to construct the null distribution of the observed variables, by centering the empirical distributions around zero so that the means of the scalar diffusion measures of WM microstructure for the controls and the 3 time points of the concussed are all zero. Then the effect at 2 days, 2 weeks and 2 months with respect to the control group are exactly zero.

Step 2. Generate a dataset with the same variables, groups and dimensions as the original data but sampled with replacement from the null distribution defined in step 1, which is also called bootstrap resampling.

Step 3. Calculate the Hoteling $T^2$ statistic from the data set generated in Step 2 and save the value.

Step 4. Repeat steps 2 and 3, 10000 times and store the 10000 values of the $T^2$ statistic. These 10000 values form the bootstrap distribution of the statistic $T^2$ under the null hypothesis.

Step 5. Calculate the bootstrap p-values by comparing the observed $T^2$ from the real data to their corresponding bootstrap distribution under the null.

References


Cubon VA, Putukian M, Boyer C, Dettwiler A. A Diffusion Tensor Imaging Study


Huisman TAGM, Schwamm LH, Schaefer PW, Koroshetz WJ, Shetty-Alva N,


