Serum Neurofilament Light in American Football Athletes over the Course of a Season

Jonathan M. Oliver¹, Margaret T. Jones², K. Michelle Kirk³, David A. Gable³, Justin T. Repshas¹, Torie A. Johnson¹, Ulf Andreasson⁵, Niklas Norgren⁵, Kaj Blennow⁶, Henrik Zetterberg⁶,⁷

¹Department of Kinesiology, Texas Christian University, Fort Worth, TX
²Division of Health and Human Performance, George Mason University, Manassas, VA
³Department of Sport Medicine, Texas Christian University, Fort Worth, TX
⁴Department of Sport Medicine, John Peter Smith Hospital, Fort Worth, TX
⁵UmanDiagnostics, Umeå, Sweden
⁶Clinical Neurochemistry Laboratory, Institute Of Neuroscience and Physiology, the Sahlgrenska Academy at University of Gothenburg, Mölndal, Sweden
⁷Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, UK

Corresponding Author:
Jonathan M. Oliver, Ph.D.
Texas Christian University
Department of Kinesiology
TCU Box 297730
Fort Worth, TX 76129
(817)257-5623 Office
(817)257-7702 Fax
jonathan.oliver@tcu.edu
Abstract

Despite being underreported, American football boasts the highest incidence of concussion among all team sports, likely due to exposure to head impacts that vary in number and magnitude over the season. This study compared a biological marker of head trauma in American football athletes with non-contact sport athletes and examined changes over the course of a season. Baseline serum neurofilament light polypeptide (NFL) was measured after 9 weeks of no contact and compared with a non-contact sport. Serum NFL was then measured over the course of the entire season at eight time-points coincident with expected changes in likelihood of increased head impacts. Data were compared between starters (n = 11) and non-starters (n = 9). Compared with non-starters (mean ± standard deviation) (7.30 ± 3.57 pg•mL⁻¹) and controls (6.75 ± 1.68 pg•mL⁻¹), serum NFL in starters (8.45 ± 5.90 pg•mL⁻¹) was higher at baseline (mean difference; ±90% confidence interval) (1.69; ± 1.96 pg•mL⁻¹ and 1.15; ± 1.4 pg•mL⁻¹, respectively). Over the course of the season, an increase (effect size [ES] = 1.8; p < 0.001) was observed post-camp relative to baseline (1.52 ± 1.18 pg•mL⁻¹), which remained elevated until conference play, when a second increase was observed (ES = 2.6; p = 0.008) over baseline (4.82 ± 2.64 pg•mL⁻¹). A lack of change in non-starters resulted in substantial differences between starters and non-starters over the course of the season. These data suggest that a season of collegiate American football is associated with elevations in serum NFL, which is indicative of axonal injury, as a result of head impacts.
The Centers for Disease Control (CDC) estimate 1.6 to 3.8 million concussions occur in sports and recreational activities each year. In the United States, American football boasts the highest number of participants and, subsequently, reports the highest incidence of concussion. Concussion awareness has increased over the years, yet a longitudinal study in professional American football athletes discovered a constant rate of concussion incidence over a recent 12 year span, suggesting increased awareness has not reduced the number of concussive incidents in American football athletes. Further, while rules and equipment have evolved over time to protect athletes better, American football players have experienced the greatest increases in body mass and size over the past decade, which may contribute to increased mechanical forces associated with the impacts in the sport resulting in more severe concussions. More importantly, athletes who have endured a concussion have a significantly greater risk of subsequent concussions.

Head trauma displays the characteristic signs of inflammation such as edema, activated resident immune cells, and impaired function. The primary insult refers to the mechanical damage that occurs at impact. The secondary injury process is the pathological response that is triggered once the injury occurs. However, despite characteristic signs of inflammation, in the absence of overt structural damage, clinical assessments remain the only tool available for diagnosing a concussive injury. While a number of those tests have been validated, they are subjective in nature and can result in misdiagnosis. Such a misdiagnosis may result in athletes returning to play before full recovery, which may subject the athlete to further injury and increase the likelihood of damage to the brain. Various biomarkers to identify brain injury have been discovered; however, to our knowledge, no study to date has examined the effect of an entire American football season on any of those biomarkers in American FB athletes.

Of the recently identified biomarkers, neurofilament light polypeptide (NFL) has been suggested to be the most sensitive. Neurofilaments, key intermediate fibers in neurons, are major components of the axonal skeleton; thus marked elevations in peripheral blood indicate some level of axonal
injury. Elevations of NFL have been reported in the cerebrospinal fluid (CSF) of boxers in the absence of a concussion diagnosis. 11-13 Like boxing, the sport of American football is physical in nature resulting in exposure to head impacts that vary in magnitude and number over the course of a season, regardless of level of play 16-23. While repeat sampling of CSF is impractical and unethical, routine blood samplings provide a potential method to assess level of head trauma. Recent technological advancements in immunoassays, like the single-molecule immunoarray (Simoa) developed by Quanterix, Inc.; 24 allow detection that is approximately 1000-fold over that of conventional assays, increasing the likelihood of identifying changes in serum NFL over the course of the season in this population. Therefore, in this study, we examined concentrations of serum NFL over the course of a season in a collegiate American football team. Further, baseline serum NFL was compared to an age matched athletic population with low risk of head trauma.

Methods
This study was an observational study from June 2014 to December 2014. All procedures involving human subjects were approved by the Institutional Review Board of Texas Christian University for use of human subjects in research (protocol no. 1404-68-1404) and were conducted according to the Declaration of Helsinki guidelines. Written consent was obtained from all subjects.

One-hundred sixteen (n = 116) Division I National Collegiate Athletic Association (NCAA) American football athletes were enrolled in the study to examine differences in baseline serum NFL. Nineteen (n = 19) Division I NCAA swim athletes served as controls. Of the 116 American football athletes, nineteen (n = 19) volunteered to provide blood samples over the course of a season. All athletes were cleared to participate in university athletics as determined by the team physician. Athletes injured or unable to participate in regularly scheduled conditioning or competitions were excluded.
**Study Design.** A between groups design was used to examine differences in serum NFL in a group of high concussion risk athletes, American football athletes, and athletes at low risk for concussion, swimmers. Further, given that some American football athletes are known to have a higher number of total contact hours due to the number of plays (repetitions) per game, baseline serum NFL was compared across three groups: starters, non-starters, and swimmers. Starters were defined as athletes known to go out with the first or second team, first or second on the depth roster, and take a majority of the repetitions (~20-40+ per game). Athletes with significant playing time (~20+ repetitions per game) or part-time starters, second on the depth roster, were also included in the starter group. A baseline blood sample was taken when volunteers returned from a period of no contact (9-weeks) (Baseline; T1). Subsequent blood samples were obtained to determine the effect of training and competition on serum NFL over 189 days of American football to include summer conditioning (57 days), fall camp (23 days), and competition (109 days). Summer conditioning consisted of strength and conditioning workouts Monday through Friday for approximately two hours per day. A second sample was obtained at the conclusion of summer training prior to preseason training camp (Pre-Camp; T2). In the time between baseline and the start of training camp no head to head contact took place, but many of the athletes elected to participate in non-supervised seven on seven practices, which involved no contact or equipment and occurred approximately twice per week for one-hour. The third sample was collected following preseason training camp (Post-Camp; T3). Camp began with three non-contact practices in shorts and helmets, one practice in helmet and shoulder pads, followed by full gear practices. The first five days practices were held once per day. Thereafter, two-a-day practices were held every other day for the remainder of camp. One day a week (Sunday) the athletes would have off from training and practice. The remaining blood samples were taken throughout the competitive season on the Monday following a Saturday game (48 hours) (T4-T8). The first two samplings occurred during Pre-Season play, whereas the last three occurred during conference play. During the competitive
season athletes underwent a one hour shorts and helmets practice on Sunday, Monday was an off day, followed by full gear practice on Tuesday, helmets and shoulder pads on Wednesday, Thursday and Friday were shorts and helmets (no full contact, incidental only). During the competition season, tackles were not taken to the ground, hit and wrap only. Additionally, athletes would perform three mandatory strength and conditioning training sessions per week. The maximum (median) number of head impacts for an NCAA American football athlete has been reported to range from 15.6 to 24 (4.8 to 7.5) and 58.5 to 86.1 (12.1 to 16.3) for practices and games, respectively. All samplings throughout the season occurred within 14 - 28 day intervals, with no sampling occurring more than 28 days from previous.

**Blood sampling and preparation.** Fasting (≥ eight hours) blood samples were collected via venipuncture from the antecubital fossa region using standard, sterile phlebotomy procedures in serum vacutainer tube with no additive (BD Diagnostics, Franklin Lakes, NJ, USA). All samples were allowed to clot then centrifuged at 2,000 g for 30 minutes at 4°C (Beckman Coulter, Allegra X-15R, Brea, CA, USA) within 30 minutes of collection. Aliquots collected via sterile pipette were immediately transferred to a pre-labeled polypropylene vials and stored at -80°C until later analysis.

**Serum Neurofilament Light (NFL).** Serum NFL levels were determined using the NF-Light kit from UmanDiagnostics (UmanDiagnostics, Umeå, Sweden), transferred onto the Simoa platform using a homebrew kit (Quanterix Corp, Boston, MA, USA). The lower limit of quantification was 1.95 pg•mL⁻¹. The analyses were performed by a board-certified laboratory technician in one round of experiments using one batch of reagents with intra-assay coefficients of variation below 10%.

**Statistical Analyses.** All data were normally distributed. For the baseline comparison, we used an independent samples t-test. The effect of starter status (starters and non-starters) and time (season) on serum NFL were estimated from a mixed model analysis of variance (ANOVA) (SPSS Version 21.0; Armonk, NY). We used a magnitude-based inference approach to inference. In the current study, all data are mechanistic. Estimates and uncertainty (90% confidence interval) for the effect on outcome
measures were derived from the ANOVA. The magnitude threshold for the smallest change was calculated as the Cohen $d$ standardized difference ($0.2 \times SD_{between}$). The probability that a contrast was at least greater than the smallest standardized difference was reported as follows: 1.0%, almost certainly not; 1.0–5%, very unlikely; 5–25%, unlikely; 25–75%, possible; 75–95%, likely; 95–99%, very likely; 99%, almost certain. In the case where the majority (>50%) of the CI lies between the threshold for substantiveness the effect was qualified trivial (negligible). Estimated mean effect sizes were qualified from small to extremely large using a modified Cohen’s $d$ where trivial, 0.0–0.2; small, 0.2–0.6; moderate, 0.6–1.2; large, 1.2–2.0; very large, >2.0; extremely large, >4.0.

**Results**

Baseline serum NFL was 6.75±1.68 pg•mL$^{-1}$, 7.30±3.57 pg•mL$^{-1}$, and 8.45±5.90 pg•mL$^{-1}$ for controls, non-starters, and starters, respectively. Relative to control, baseline serum NFL was likely greater by a substantial moderate magnitude (ES = 0.4; $p = 0.153$) in those athletes categorized as starters (standardized difference; 90% CL) (1.69; ±1.96 pg•mL$^{-1}$). Further, a possible difference (ES = 0.3; $p = 0.176$) was observed when comparing starters and non-starters (1.15; ±1.40 pg•mL$^{-1}$). All other contrasts were unclear.

Serum NFL increased substantially over the course of the season in those participants categorized as starters (Figure 1). The observed increase occurred coincident with changes in intensity and hours of contact. The large increase (ES = 1.8; $p < 0.001$) observed Post-Camp (T3), coincident with the greatest number of contact hours, was most likely most likely greater relative to baseline (1.52; ±1.18 pg•mL$^{-1}$). Similar, albeit smaller and of a less magnitude, increase in serum NFL was observed in non-starter during the same period, as athletes in this category would have also experienced the most contact during the time from Pre-Camp (T2) to Post-Camp (T3). All other contrasts in the non-starters were unclear. However, serum NFL remained elevated the remainder of the competition season in starters, with a further very large substantial increase occurring at T6 (ES = 2.6; $p = 0.008$), which
occurred coincident with conference play. Likewise, a further increase was noted at T7 (ES = 3.2; p = 0.002), which remained elevated to the end of the competitive season (ES = 3.5; p = 0.007). The increase observed over the course of the study in starters resulted in substantial differences noted over the course of the study when compared to non-starters (Table 1). Similarly, contrasts based on determination of area under the curve resulted in a very likely substantial large magnitude difference (601; ± 430 pg/mL•day⁻¹; ES = 1.2; p = 0.026) between starters (1780 ± 629 pg/mL•day⁻¹) and non-starters (1179 ± 341 pg/mL•day⁻¹).

Discussion

Even in the presence of underreporting,⁹,²⁷ American football is the sport associated with the highest incidence of concussion². Given that a second impact while symptomatic from an initial injury can result in major trauma,²⁸,²⁹ a sensitive and reliable biological marker may provide a viable option for early detection and proper diagnosis. Recently, a number of biological markers have been identified that may prove to be reliable in concussion diagnosis¹⁰. However, to our knowledge, no study to date has reported on the effect of a season of American football on any biomarker of head trauma. The major finding of this study was that serum levels of NFL increased over the course of a season in American football athletes, specifically in those athletes categorized as starters. The increases observed in starters, which ranged in magnitude from moderate to very large following periods in which the number of head impacts may have increased, likely represents sub-concussive impacts with release of NFL from the neurons. Further, a 9-week period of no contact was not sufficient to normalize serum NFL to that observed in controls or non-starters.

Elevations in CSF NFL have been reported in amateur boxers post-bout in the absence of a clinical diagnosis of concussion.¹¹,¹³ Herein, we report that a likely substantial elevation of moderate magnitude was present at baseline in those athletes categorized as starters following the 9-week period of no contact. These data are in agreement with reports of higher CSF NFL in boxers receiving many hits
compared to those receiving fewer. However, following a 3-month rest period of no contact, no differences were detected in CSF NFL between boxers and non-boxing controls. The fact that elevations in serum NFL were present in the starters of the current study following a 9-week reprieve suggests the impacts over the course of a season in American football may be greater in magnitude than those experienced in boxing. Indeed, though boxers deliver punches with high impact velocity, a lower head injury criterion due to shorter duration acceleration and translational acceleration has been reported when compared to American football.

The average number and magnitude of head impacts varies over the course of an American football season. In the current study, the most pronounced changes observed in serum NFL occurred coincident with those times in which an increase in head impacts was likely. These data suggest that some level of damage occurs in American football athletes over the course of a season in the absence of a concussion diagnosis. NFL has been suggested to be one of the most sensitive biomarkers of head trauma, specifically axonal injury, to date. DAI typically results from head rotational acceleration/deceleration or propagation of force through the brain following impact, routine occurrences in the sport of American football. Risk weighted exposure (RWE), a metric defined as collected risk of concussion over the course of a season, shows strong associations with diffusion tensor imaging (DTI) scalars over the course of a season in high school American football athletes. In contrast to conventional neuroimaging methods, DTI allows for the identification of subtle changes in white matter enabling tracking of small changes longitudinally in patients/subjects. Further, most studies using DTI in TBI and mTBI report that the changes in those scalars associated with RWE result from demyelination or disruption of tissue structure. Given that collegiate players sustain high-level impacts more frequently than high school players, the current data are of interest and further suggest data reported herein indicate the elevations in NFL represent some level of sub-concussive impacts sustained by collegiate American football players and reflect some degree of axonal injury.
Inference to the current data is constrained by a number of limitations. The sample size of the current study was small. Further, a lack of biomechanical, neuroimaging, and neuropsychological data limit our ability to determine if the elevations in serum NFL are a result of axonal damage caused by head impacts or from another source, such as muscle. However, despite participating in practice over the course of the season, serum NFL did not vary in those categorized as non-starters, suggesting that the elevations observed herein were in fact a result of additional impact to the head. A relatively weak, but significant relationship has been reported between CSF NFL and serum NFL\textsuperscript{41}, still further studies are needed to determine the strength of that relationship.

In conclusion, our study suggests that collegiate American football athletes suffer from some level of axonal injury over the course of a season. The level of injury is likely similar, albeit of a lesser magnitude, in high school American football athletes\textsuperscript{23}, while of greater magnitude at the professional level. Future studies examining NFL and other potential biomarkers in combination with neuroimaging and neuropsychological tests are warranted in order to identify the potential effect in American football due to the high risk associated with this population.


Acknowledgements:

Dr. Oliver and Dr. Zetterberg had full access to all the data in the study and take full responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Oliver, Jones, Kirk, Gable, Repshas, Blennow, Zetterberg

Acquisition of data: Oliver, Jones, Kirk, Gable, Repshas, Johnson, Blennow, Zetterberg

Drafting of the manuscript: Oliver, Jones, Repshas, Blennow, Zetterberg

Critical revision of the manuscript for important intellectual content: Oliver, Jones, Blennow, Zetterberg

Obtained funding:

Administrative, technical, or material support: Oliver, Jones, Kirk, Gable, Blennow, Zetterberg

Study supervision: Oliver, Jones, Kirk, Gable, Repshas, Johnson, Blennow, Zetterberg

Conflict of Interest Disclosures:

Funding/Support: Texas Christian University

Role of the Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.
Table 1. Neurofilament light (pg•mL\(^{-1}\)) at each time point in STARTERS and NON-STARTERS over the duration of the study with corresponding statistical summary on effect.

<table>
<thead>
<tr>
<th>Time Point</th>
<th>Non-Starters (^a) (n = 9)</th>
<th>Starters (^a) (n = 11)</th>
<th>Mean Difference(^bc)</th>
<th>Inference(^d)</th>
<th>Effect Size(^e) (Magnitude)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (T1)</td>
<td>5.69±1.18</td>
<td>6.50±1.85</td>
<td>0.81; ±1.48</td>
<td>Unclear</td>
<td>Small</td>
</tr>
<tr>
<td>Pre-Camp (T2)</td>
<td>6.49±1.93</td>
<td>7.65±3.02</td>
<td>1.17; ±2.13</td>
<td>Unclear</td>
<td>Small</td>
</tr>
<tr>
<td>Post-Camp (T3)</td>
<td>6.70±2.03</td>
<td>9.84±3.69</td>
<td>3.14; ±2.52</td>
<td>Very likely</td>
<td>Moderate</td>
</tr>
<tr>
<td>Competition (T4)</td>
<td>6.12±2.42</td>
<td>8.34±2.63</td>
<td>2.22; ±2.05</td>
<td>Likely</td>
<td>Moderate</td>
</tr>
<tr>
<td>Competition (T5)</td>
<td>6.06±2.14</td>
<td>8.70±3.22</td>
<td>2.63; ±2.28</td>
<td>Likely</td>
<td>Moderate</td>
</tr>
<tr>
<td>Competition (T6)</td>
<td>5.95±2.85</td>
<td>11.32±7.92</td>
<td>5.37; ±5.12</td>
<td>Likely</td>
<td>Moderate</td>
</tr>
<tr>
<td>Competition (T7)</td>
<td>6.52±1.78</td>
<td>12.47±6.88</td>
<td>5.95; ±4.37</td>
<td>Very likely</td>
<td>Large</td>
</tr>
<tr>
<td>Competition (T8)</td>
<td>6.05±1.57</td>
<td>13.00±8.85</td>
<td>6.95; ±5.56</td>
<td>Most likely</td>
<td>Large</td>
</tr>
</tbody>
</table>

\(^a\) Serum neurofilament light (pg•mL\(^{-1}\)) with corresponding SD over the course of the season in American football athletes.

\(^b\) Mean difference between Starters and Non-Starters.

\(^c\) 90% CL; Add or subtract from the mean difference to obtain the upper and lower confidence limits.

\(^d\) Magnitude-based inference about the true value for outcomes where the threshold for smallest substantial change was calculated as 0.2 times baseline between-subjects SD.

\(^e\) Qualified according to a modified Cohen’s \(d\) effect size where trivial, 0.0-0.2; small, 0.2-0.6; moderate, 0.6-1.2; large, 1.2-2.0; very large, >2.0; extremely large, >4.0.
Figure Legends

Figure 1. Serum neurofilament light (pg•mL⁻¹) over the course of the study in starters and non-starters; Threshold for smallest substantial change calculated as 0.2 times baseline SD between; Likelihood is shown as increased number of symbols (+) relative to baseline in each treatment group: + possible, ++ likely, +++ very likely, ++++ most likely. All data are mean. A magnitude-based summary of statistical contrasts between starters and non-starters with corresponding mean ± SD is shown in Table 3.