Evaluation of an In-ear Sensor System for Quantifying Head Impacts in Youth Football (Soccer)

ABSTRACT

Background: Wearable sensor systems have the potential to quantify head kinematic responses of head impacts in soccer. However, on-field use of sensors, e.g. accelerometers, remains challenging, due to poor coupling to the head and difficulties discriminating low-severity direct head impacts from inertial loading of the head from human movements, such as jumping and landing.

Purpose: To test the validity of an in-ear sensor for quantifying head impacts in youth soccer.

Study design: Descriptive laboratory and on-field study.

Methods: First, the sensor was mounted to a Hybrid III headform (HIII) and impacted with a linear impactor or a soccer ball. Peak linear acceleration (PLA), peak rotational acceleration (PRA) and peak rotational velocity (PRV) were obtained from both systems; random and systematic error were calculated using HIII as reference. Then, six youth soccer players wore sensors and performed a structured training protocol including heading and non-heading exercises; they also completed two regular soccer sessions. For each accelerative event recorded, PLA, PRA and PRV outputs were compared to video recordings. Receiver operating characteristic curves were used to determine the sensor's discriminatory capacity in both on-field settings, determining cut-off values for predicting outcomes.

Results: For the laboratory tests, the random error was 11% for PLA, 20% for PRA and 5% for PRV; the systematic error was 11%, 19% and 5%, respectively. For the structured training protocol, heading events resulted in higher absolute values (PLA=15.6±11.8g) than non-heading events (PLA=4.6±1.2g); the area under the curve (AUC) was 0.98 for PLA. In regular training sessions, AUC was >0.99 for PLA. A 9g cut-off value yielded a positive
predictive value of 100% in the structured training protocol vs. 65% in regular soccer sessions.

Conclusion: The in-ear sensor displayed considerable random error and overestimated head impact exposure substantially. Despite showing an excellent on-field accuracy for discriminating headings from other accelerative events in youth soccer, absolute values must be interpreted with caution, and there is a need for secondary means of verification (e.g. video analysis) in real-life settings.

Clinical Relevance: Wearable sensor systems can potentially provide valuable insights into head impact exposures in contact sports, but their limitations require careful consideration.

Key words: TBI, REPEETITIVE, SOCCER, SUBCONCUSSIVE, WEARABLE, ACCELEROMETER
INTRODUCTION

Repetitive head impacts in the ‘subconcussive’ range (i.e. head impacts without immediate symptoms) are common in soccer where purposeful and unprotected heading of the ball is an integral part of the game. There is evidence of long-term brain structural and functional alterations in soccer players.\textsuperscript{5,6,8,9} Moreover, recent studies suggest a potential effect on cognitive function in children and adolescents during a vulnerable period of brain maturation.\textsuperscript{7,22} However, the link between exposure to repetitive head impacts and brain alterations is still controversial and remains to be elucidated. In this context, accurate measurement of head impact exposure in soccer is a key challenge when investigating the effect of head impact exposure on brain health.

Wearable sensor systems, such as accelerometers/gyroscopes, can potentially provide valuable insights into the dynamics of single and repetitive head impacts. However, several issues have made quantifying head impact exposure challenging, despite the multiple systems currently available.\textsuperscript{12,13} A central issue has been poor sensor coupling with the head; methods such as skin patches and skull caps are subject to relative motion between the device and the skull, and therefore not able to measure head impact exposure \textit{in vivo} accurately.\textsuperscript{20} This issue extends beyond erroneous outputs for direct head impacts; failure to discriminate these from non-head impact accelerative events typically seen during game play (running, jumping, tackling etc.) also leads to high false positive rates.\textsuperscript{3,14} Thus, previous studies have concluded that secondary means of verification, such as video analysis, are needed to verify whether the events recorded actually represent head impacts.\textsuperscript{2,3,14} This makes surveillance of exposure in large-scale cohort studies considerably more demanding.

Recently, in-ear sensor systems have become commercially available, aiming to improve skull coupling by custom-molded placement in the bony portion of the external ear canal. However, before usage in prospective cohort studies, it is necessary to evaluate their
performance in both a laboratory and an on-field setting. The aim of this paper was to test the
validity of a new in-ear sensor for quantifying head impacts in youth soccer.

METHODS

Study design and participants
This study was conducted in three separate phases: (1) validation of the in-ear sensor in a
controlled laboratory setting, (2) controlled on-field evaluation of its ability to differentiate
headings from other accelerative events typically seen in soccer, and, finally, (3) on-field
evaluation in regular soccer training sessions. In phases 2 and 3, six male youth soccer players
(age 15.3±0.3 years, height 170.3±5.0 cm, mass 54.8±6.1 kg) participated, all playing at the
regional elite youth level in Norway during the 2017 season. The Ethics committee at the X
institution approved the study, and written informed consent was obtained from the
participants and their legal guardians.

The in-ear sensor
MV1 (MV1, MVTrak, Durham, NC, USA) is a sensor system designed for custom-molded
placement in the left external ear canal to optimize coupling to the head. A small lumen runs
through the sensor to allow air conduction, limiting hearing loss to approximately 3 dB. The
sensor samples linear acceleration and rotational velocity data at 1000 Hz, filtering the data
with a phaseless 300 Hz 8-pole low-pass Butterworth filter to remove noise; rotational
acceleration is calculated by differentiating this filtered rotational velocity data. The sensor
then provides a time-stamped output of peak linear acceleration (PLA), peak rotational
velocity (PRV) and peak rotational acceleration (PRA) for all accelerative events exceeding
3 g (i.e. nominal head impacts), followed by a 250 ms latency period before another impact
can be registered. The sensor stores event-specific data on a microchip, and connects via USB
to a computer for download. Raw data are uploaded to the MVTrak server, before being
processed by the producer's algorithm. These data can then be downloaded for each player
(i.e. sensor) as time-stamped and event-specific summaries in Excel charts, including PLA,
PRV, PRA and the individual kinematic components of each accelerative event.

Experimental procedures

Phase 1, Laboratory validation. The MV1 sensor was mounted at the ear region of an in-
calibration Hybrid III (HIII) head and neck assembly. Three mounting configurations were
assessed: (1) a custom-made flat MV1 sensor (MV1 flat) attached with double-sided tape,
reinforced with external taping to minimize relative motion between the HIII headform and
the sensor, to optimize the coupling conditions and assess this as an alternative to in-ear
mounting; (2) a regular in-ear MV1 (MV1 in-ear) firmly placed in a tight canal on the HIII
headform, representing an artificial ear canal; and (3) a regular in-ear MV1 (MV1 loose)
loosely placed by expanding the same canal (figure 1). We created the canal by carving out a
piece of the artificial skin covering of the HIII headform. The tight canal's diameter was
slightly smaller than the sensor's, only enough to allow the compliant properties of the rubber
to expand and create a snug fit, mimicking real-life custom-molded placement; the expanded
canal's diameter was slightly larger (2-3 mm) than the sensor's, allowing slight relative motion
for the sensor. The HIII head was instrumented at its centre of mass with an in-calibration
triaxial linear accelerometer and triaxial angular velocity sensor array sampling at a rate of 20
kHz. Linear acceleration and angular velocity data were filtered with a SAE CC1000 filter
and a SAE CC180 filter, respectively, before computing a preliminary set of PLA and PRV
values for each impact. PRA values were calculated by differentiating the filtered angular
velocity data. HIII-measured impact characteristics were considered reference values for MV1
flat; for evaluating between-sensor agreement, MV1 flat was considered reference for MV1
in-ear. The HIII was impacted at selected locations over a range of selected magnitudes with (1) a linear impactor with a stiff interface, (2) a linear impactor with a compliant interface, or (3) a FIFA-approved soccer ball inflated to 11 PSI (table 1). Each test was videoed with a GoPro HERO5 Black camera, recording at 240 Hz.

Figure 1. Mounting of the MV1 in-ear (left) and MV1 flat (left middle) on the Hybrid III headform, and an example of a setup for right frontal impacts with the padded impactor striking from a 45 degree angle (right middle). Shown to the right is a photo of the MV1 in a real-life setup.

Phase 2, Controlled on-field evaluation. The six participants were invited to complete a structured training protocol in a controlled setting twice while wearing a custom-molded MV1 in their left ear canal. The protocol was designed and supervised by research staff with long-standing experience in soccer (author X and author Y), and consisted of five heading and six non-heading exercise drills typical for soccer. Heading exercises included (1) finishing headers, (2) redirectional headers, (3) long direct headers, (4) short direct headers, and (5) headers from in-air duels. Non-heading exercises included (1) shoulder-to-shoulder collisions, (2) forceful shooting, (3) redirectional running with maximal intensity, (4) short straight sprinting with maximal intensity, (5) falling abruptly forward on the ground landing on outstretched arms, and (6) in-air duels without ball contact (losing the duel).
Phase 3, In-training on-field evaluation. The participants wore the sensors for two regular training sessions with their team. The sessions were instructed by their regular coaching staff, and included warm-up, passing and playing drills, in addition to regular play in teams.

Phases 2 and 3 were performed on artificial turf in an outdoor setting. Video recordings were obtained with two digital video cameras (1080p, 50 fps), placed to capture all movements on the pitch to subsequently verify and classify events.

Data analyses

For the laboratory validation, the HIII kinematic time histories (e.g. linear acceleration) were reviewed comparing with high-speed video of each test. The aim was to review the preliminary PLA, PRA and PRV values for each test and to identify the peak values directly related to the initial interaction between the impactor/soccer ball and the HIII headform. After review by authors X and Y, a final set of HIII PLA, PRV and PRA values was determined.

To estimate the accuracy of the MV1 sensor for different impact types, locations and mounting configurations, we calculated its random and systematic error. The random error was calculated by first dividing the standard deviation (SD) of the mean difference between the MV1 and the reference (HIII) by the square root of the number of measurements (n=2); this value was then divided by the mean of the combined measurements, expressing the random error as a percentage.\textsuperscript{16} The systematic error was calculated as the mean difference between the sensor and the reference, divided by the mean reference value. Expressed as a percentage, positive and negative results indicate systematic overestimation and underestimation by the MV1, respectively. For the soccer ball impacts, MV1 flat and MV1 in-ear were mounted to the HIII simultaneously; agreement between the two sensors were expressed with the same formulas, using MV1 flat as the reference.
For the structured training protocol (phase 2), the individual events of each exercise drill were used as reference and compared to the time-stamped outputs from the sensors. If an event failed to exceed the sensor's 3 g threshold, and therefore was not recorded, kinematic values were set as follows to be included in later analyses: PLA=3.0 g, PRV=3.0 rad/s, and PRA=200 rad/s². These values were set arbitrarily, assuming that these events involved slightly lower values than the lowest magnitude events recorded from the sensor; this was done to include them in the ROC analyses. For the regular training sessions (phase 3), all head impacts were first identified on video to be included in the analyses; they were then compared to their potential time-stamped sensor outputs. All other nominal head impact events recorded by the sensors (i.e. either non-head impact accelerative events or spurious events) were then classified according to video.

For both on-field evaluations (phases 2 and 3), mean values ± SD for PLA, PRV and PRA were calculated for (1) all head impact and (2) all non-head impact events; this was done separately for the structured training protocol and regular training sessions. To test if head impacts resulted in higher absolute peak values, independent-samples t-tests were used to compare the means of the two event groups in both settings, using an a priori significance level of p<0.05. Then, to determine the discriminatory capacity, receiver operating characteristic (ROC) curves were constructed for each dependent variable (PLA, PRA and PRV) in both settings. Expressed as area under the curve (AUC), results were interpreted as excellent (1.0-0.9), good (0.9-0.8), fair (0.8-0.7) or fail (0.7-0.6). To investigate how the sensor would perform in settings without other verification means, sensitivity and positive predictive value were then calculated in both settings according to different PLA or PRV cut-off values identified from the ROC curve.

SPSS version 24 (IBM SPSS Statistics, IBM Corporation, Chicago, IL) was used for all statistical analyses.
RESULTS

**Phase 1, Laboratory validation.** For MV1 flat, 112 impacts were included for final analyses (table 1). When reviewing HIII outputs, we excluded angular kinematic data only (i.e. PRA and PRV) from one of the 112 impacts, since we were unable to identify the appropriate initial peak values. Furthermore, for one series of consecutive impacts (n=12), all within the same time period with identical set-up on the same afternoon, angular kinematic values (PRA and PRV) from MV1 flat were recognized as severe outliers (values ranging from four to 13 times higher than the reference). Upon our request, the MV1 producer reviewed the data for these specific impacts, and suspected that vibrations between the MV1 flat and the HIII was the cause. We replaced these data points with outputs from MV1 in-ear from the same impacts.

As shown in figure 2, PLA values showed the strongest correlation, followed by PRV and PRA. The random error for all impacts was 11% for PLA, 20% for PRA and 5% for PRV. The systematic error was 11% for PLA, 19% for PRA and 5% for PRV. The random error varied with impact type and location, consistently overestimating PLA, PRA and PRV values (table 1). When testing for agreement between MV1 flat and MV1 in-ear for the soccer ball impacts (n=29 for PLA; n=28 for PRA and PRV values), using MV1 flat as reference, the random error was 6% for PLA, 20% for PRA and 6% for PRV; the systematic error was -5% for PLA, -23% for PRA and -3% for PRV.

For MV1 loose, we replicated seven right frontal impacts and one frontal impact (HIII PLA range: 29-122 g) also used for mounting configuration 2 (i.e. MV1 in-ear). Compared to MV1 in-ear, the loose coupling in mounting configuration 3 led to an increase in the random error from 10% to 14% for PLA, 10% to 55% for PRA, and 7% to 20% for PRV. Systematic error increased from 17% to 33% for PLA, 19% to 202% for PRA, and 13% to 32% for PRV.
TABLE 1. Comparison between the reference (Hybrid III headform) and MV1 flat, with random and systematic error of PLA, PRA and PRV values, according to impact type and location.

<table>
<thead>
<tr>
<th>Impact type and location</th>
<th>No. of impacts</th>
<th>PLA range (g)</th>
<th>PRA range (rad/s²)</th>
<th>PRV range (rad/s)</th>
<th>PLA (g) Random error (%)</th>
<th>Systematic error (%)</th>
<th>PRA (rad/s²) Random error (%)</th>
<th>Systematic error (%)</th>
<th>PRV (rad/s) Random error (%)</th>
<th>Systematic error (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear impactor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>37</td>
<td>26 - 132</td>
<td>1121 - 6901</td>
<td>12 - 23</td>
<td>3</td>
<td>4</td>
<td>14</td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Right frontal</td>
<td>21</td>
<td>27 - 110</td>
<td>1755 - 8030</td>
<td>12 - 20</td>
<td>9</td>
<td>28</td>
<td>18</td>
<td>21</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Right zygomatic</td>
<td>12</td>
<td>27 - 138</td>
<td>1835 - 5087</td>
<td>16 - 26</td>
<td>5</td>
<td>6</td>
<td>11</td>
<td>45</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Right temple</td>
<td>13</td>
<td>25 - 144</td>
<td>1668 - 11537</td>
<td>11 - 20</td>
<td>12</td>
<td>-4</td>
<td>13</td>
<td>-6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>25 - 144</td>
<td>1121 - 11537</td>
<td>11 - 26</td>
<td>10</td>
<td>8</td>
<td>18</td>
<td>15</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Soccer ball</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal</td>
<td>9</td>
<td>9 - 20</td>
<td>997 - 2203</td>
<td>5 - 11</td>
<td>17</td>
<td>33</td>
<td>30</td>
<td>54</td>
<td>2</td>
<td>-1</td>
</tr>
<tr>
<td>Right frontal</td>
<td>7</td>
<td>13 - 22</td>
<td>958 - 4638</td>
<td>7 - 13</td>
<td>16</td>
<td>67</td>
<td>29</td>
<td>40</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>Frontal/crown</td>
<td>10*</td>
<td>13 - 26</td>
<td>1362 - 3343</td>
<td>7 - 14</td>
<td>17</td>
<td>39</td>
<td>38</td>
<td>39</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Face</td>
<td>3</td>
<td>11 - 19</td>
<td>722 - 3352</td>
<td>6 - 10</td>
<td>15</td>
<td>40</td>
<td>18</td>
<td>6</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>29*</td>
<td>9 - 26</td>
<td>722 - 4638</td>
<td>5 - 14</td>
<td>18</td>
<td>45</td>
<td>33</td>
<td>39</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>

PLA, peak linear acceleration. PRA, peak rotational acceleration. PRV, peak rotational velocity.
*PRA and PRV values were excluded for one impact.
Figure 2. Peak linear acceleration (A), peak rotational acceleration (B) and peak rotational velocity (C) from MV1 flat plotted against the reference (Hybrid III headform). Linear regression lines (dotted) with reference lines (solid) are for all head impacts combined (i.e. with linear impactor and soccer ball).
Phase 2, controlled on-field evaluation. All six participants completed each exercise drill at least once, with the number of events obtained per drill ranging from 44 to 180. Heading events (n=431) resulted in higher average values for all three variables (PLA=15.6±11.8 g, p<0.001; PRA=10543±10854 rad/s², p<0.001; PRV=35.1±18.3 rad/s, p<0.001) compared to non-heading events (n=750) (PLA=4.6±1.2 g; PRA=1095±823 rad/s²; PRV=9.8±4.6 rad/s). ROC curve analyses revealed an AUC of 0.98 (95% CI 0.98 to 0.99, p<0.001) for PLA, 0.99 (95% CI 0.99 to 1.00, p<0.001) for PRA and 0.97 (95% CI 0.96 to 0.98, p<0.001) for PRV. Figure 3 shows the distribution of peak values for each specific exercise.

Phase 3, in-training on-field evaluation. Five of the participants completed one or both of the regular training sessions, and, from the resulting eight sessions, the MV1 sensors recorded 2 039 nominal head impact events. Of these, 15 events were confirmed on video analysis to be direct head impacts (PLA=20.7±10.6 g, p<0.001; PRA=14541±7994 rad/s², p<0.001; PRV=43.5±16.4 rad/s, p<0.001), all of them due to purposeful heading of the ball. No other head impacts were identified on video. The remaining 2 024 events were triggered by non-head impact events such as jumping, tackling, running with change of direction, touching or losing the sensor (PLA=4.0±3.1 g; PRA=835±2541 rad/s²; PRV=7.4±4.9 rad/s), resulting in an AUC of >0.99 (95% CI 0.99 to 1.00, p<0.001) for both PLA, PRA and PRV. Tables 2 and 3 show sensitivity and positive predictive value for different cut-off values for PLA and PRV, for both the structured training protocol and the regular training sessions.
Figure 3. Box plots showing median value and interquartile range of peak linear acceleration, peak rotational acceleration and peak rotational velocity from MV1 for the exercises from the structured training protocol. The left and right markers are for the 5th and 95th percentile, respectively.
TABLE 2. MV1 sensitivity and positive predictive value for classifying accelerative events as head impacts (i.e. headers) or non-head impacts for different peak linear acceleration (g) cut-off values.

<table>
<thead>
<tr>
<th>Cut-off value (g)</th>
<th>Sensitivity (%)</th>
<th>Positive predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training protocol</td>
<td>Regular training</td>
</tr>
<tr>
<td>&gt;6</td>
<td>96</td>
<td>100</td>
</tr>
<tr>
<td>&gt;7</td>
<td>90</td>
<td>93</td>
</tr>
<tr>
<td>&gt;8</td>
<td>83</td>
<td>87</td>
</tr>
<tr>
<td>&gt;9</td>
<td>73</td>
<td>87</td>
</tr>
<tr>
<td>&gt;10</td>
<td>65</td>
<td>87</td>
</tr>
</tbody>
</table>

TABLE 3. MV1 sensitivity and positive predictive value for classifying events as head impacts (i.e. headers) or non-head impacts for different peak rotational velocity (rad/s) cut-off values.

<table>
<thead>
<tr>
<th>Cut-off value (rad/s)</th>
<th>Sensitivity (%)</th>
<th>Positive predictive value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training protocol</td>
<td>Regular training</td>
</tr>
<tr>
<td>&gt;10</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>&gt;15</td>
<td>92</td>
<td>100</td>
</tr>
<tr>
<td>&gt;20</td>
<td>75</td>
<td>93</td>
</tr>
<tr>
<td>&gt;25</td>
<td>61</td>
<td>93</td>
</tr>
<tr>
<td>&gt;30</td>
<td>52</td>
<td>80</td>
</tr>
</tbody>
</table>

DISCUSSION

This is the first study to investigate the validity of using in-ear sensors to quantify head impact exposures in youth soccer. We found that the sensor systematically overestimated head kinematic parameters and with a considerable random error (phase 1). Still, the accuracy for discriminating headers from non-head impact accelerative events in a controlled on-field setting was excellent (phase 2). However, as the proportion of head impacts (i.e. headers) was relatively low compared to non-head impact events, false positive results nevertheless remained a challenge when used in the real-life setting (phase 3).
Obtaining accurate results from compact, wearable sensor systems is difficult, as shown by Cummiskey et al.\textsuperscript{4} and others\textsuperscript{11,18,19}. A recent review by Patton\textsuperscript{13} described multiple examples of large discrepancies even in controlled laboratory settings. In the laboratory validation (phase 1), we therefore aimed to test the technical performance of the in-ear sensor, optimizing all factors, including coupling to the head. We found a consistent systematic overestimation for all peak values (PLA, PRA and PRV) and with a considerable random error, varying with impact type and location. Even though the exact reasons for this are uncertain, several previously recognized technical limitations such as low sampling rate (1 kHz for the in-ear sensor vs. 20 kHz for the reference system) might account for some of the discrepancy. The observation that the PRA component generally performed poorer than PLA and PRV simply reflects that it is derived from PRV, rendering it more susceptible to noise and to the relatively low sample rate. This is consistent with the finding that PRA values also displayed considerably poorer agreement between sensors (approx. 80%), compared to both PLA and PRV (approx. 95%). As an additional barrier, algorithms of any externally mounted system need to correct for its relative position on the head, in order to measure what is happening at the center of mass.

As we were interested in how on-field conditions could affect sensor performance in phases 2 and 3, we included a loose mounting configuration in phase 1. The idea was to test how poor coupling could affect the inherent issues described above. With an unfavorable effect on both systematic and random error for all variables, we observed a ten-fold increase in the systematic error for PRA. We believe this effectively illustrates why one should interpret absolute kinematic values from sensor systems in contact sports with caution. We suspect that the main explanation for some of the very high on-field values observed (see figure 3) is a combination of inherent systematic overestimation and poor head coupling. Arguably, a mean value of well over 20 krad/s\textsuperscript{2} for finishing headers almost certainly
represents a gross overestimation, based on previous biomechanical studies from heading in soccer and mild traumatic brain injuries\(^1,10,17,21\); the players considered the exercise to be in the upper but normal heading severity range.

Press and Rowson\(^14\) recently quantified head impact exposure in collegiate women’s soccer using a skin patch placed behind the ear. They observed that the recorded number of head impacts vastly exceeded those confirmed on video, concluding that data from head impact sensors warrant careful interpretation when used in automated settings. Cortes et al.\(^3\) drew similar conclusions when measuring head impact exposure in lacrosse, both studies highlighting the need to classify accelerative events with e.g. video analysis\(^3,14\). Thus, the main objective of the structured training protocol (phase 2) was to evaluate the in-ear sensor’s capacity to discriminate head impacts from non-head impact accelerative events. Classifying all recorded accelerative events into these two main categories, in both the structured training protocol and the regular training sessions, our results showed that the sensor displayed an excellent discriminatory capacity. However, despite the ability to maintain high sensitivity and specificity, there is a crucial difference between the two on-field settings, with real-life implications. In the structured training protocol, it was possible to use a cut-off value (e.g. 9 g, see table 2) yielding 100% positive predictive value, while still maintaining a sensitivity over 70%. In such a scenario, although missing many head impacts in the lowest range, one can safely conclude that any event above this threshold is actually caused by a direct impact to the head, obviating secondary means of verification (e.g. video). We were unable to replicate this finding in the regular training sessions (phase 3) due to spurious non-head impact events, such as touching or dropping the sensor on the ground, recording values as high as 65 and 124 g. Tables 2 and 3 illustrate the difficulties of identifying a PLA or PRV cut-off value in a real-life setting, and how it is not possible to maximize the positive predictive value in a similar manner as for the structured training protocol. Thus, there is still a need to confirm what
actually caused any event above a given threshold. During the regular training sessions we
observed, headers were relatively infrequent. But even if a greater proportion of headers most
likely would yield higher positive predictive values, there would still be a need for e.g. video
confirmation.

As the main aim of this study was to evaluate the sensor's potential for usage in large-
scale data collection in youth soccer, practical considerations on feasibility and user-
friendliness also need to be addressed. We encountered several software problems during the
course of the study, such as having to retrieve apparently missing data from one of the
on-field sessions. Furthermore, player opinion differed as to whether or not they would accept
wearing the sensors over longer periods throughout the season, including matches. Despite
being designed with a lumen to minimize any hearing impairment, this seemed to be one of
the main criticisms. We also observed that some of the sensors were partially obstructed with
cerumen after the sessions. Such concerns are likely to limit the utility of such devices; they
not only render the data potentially unreliable, but might also negatively affect compliance.

We acknowledge several study limitations. First, a laboratory validation needs to rely on
a reference system, with its own imperfections. We chose a well-recognized method (HIII) to
make our results comparable to the work of others, as well as easy to replicate. Initially, we
performed a thorough assessment of frontal impacts (considered most relevant for soccer),
then proceeding to address the issues of impact location and severity. This explains the
discrepant number of impacts across conditions. We chose to exclude and replace data from a
series of severe outliers. We did this as we consider the suspected cause plausible: a specific
mechanical response of the HIII head and neck during a sequence of impacts gave rise to an
artefact in the MV1 sensor. Such an artefact may reflect specific technical sensor
characteristics, including sample rate and sensor resonant frequency response or bandwidth.
Including these data would potentially disguise our main findings, as this particular issue does
not reflect a challenge related to the real-life human scenarios we ultimately evaluated.

Second, we recognize that only six players took part in this study and that only two regular training sessions were included, potentially limiting the external validity to other playing levels, sex, and styles of play. Compensating for this, we have a data set comprised of several hundreds of events. Last, due to logistical reasons, we attached the sensors ourselves for the on-field parts of the study, without an on-site demonstration recommended by the producer. Even though this might also be a source of systematic error, we did our best to comply with their instructions. In summary, however, it seems unlikely that these limitations invalidate our main findings.

The main strength of this study lies in its stepwise approach, allowing us to translate our findings from the laboratory into a real-life setting. As a result, we believe our findings have illustrated several challenges that needs to be taken into account when considering using such sensor systems for quantifying head impact exposures in any collision or contact sport. We suggest that new methods are evaluated carefully before taken into use, including not only a laboratory validation, but also an on-field evaluation. Future sensor systems should seek to improve technical specifications (e.g. sampling rate), create algorithms better capable of filtering out spurious non-head impact events, and optimize head coupling. Until then, it is important to remain critical when interpreting data acquired from such systems and to confirm all events with secondary means of verification.
CONCLUSION

This study highlights several previously recognized challenges when attempting to quantify head impacts in contact sports with sensor systems. It also demonstrates the need for careful and systematic evaluation before being used in real-life and research settings. In-ear sensors represent a novel method for quantifying head impact characteristics in youth soccer.

However, the device tested in this study displayed considerable random error and overestimated head impact exposures substantially, depending on both the severity and type of impact. Despite showing an excellent on-field accuracy for discriminating headings from other accelerative events in youth soccer, absolute values should be interpreted with caution, and there is a need for secondary means of verification (e.g. video analysis) in real-life settings.
REFERENCES


