- 1 Preferred Running Head: Microsensors and rugby scrums 2
- 3 *Title:* Validity of a microsensor-based algorithm for detecting scrum events in rugby union
- 4 5

7

9

11

12 13

14

15

- 6 Submission Type: Original Investigation
- 8 *Authors:* Ryan M. Chambers^{1,2}, Tim J. Gabbett^{3,4}, Michael H. Cole²
- 10 Institutions and Affiliations:
 - 1. Welsh Rugby Union, Cardiff, United Kingdom
 - 2. School of Exercise Science, Australian Catholic University, Brisbane, Queensland, Australia.
 - 3. Gabbett Performance Solutions, Brisbane, Queensland, Australia
 - 4. University of Southern Queensland, Institute for Resilient Regions, Ipswich, Queensland, Australia
- 16 17
- 18 Address correspondence:
- 19 Mr. Ryan Chambers
- 20 Welsh Rugby Union,
- 21 Westgate Street, Cardiff, UK
- 22 Email: rchambers@wru.wales
- 23
- 24 Abstract word count: 247
- 25 Text only word count: 3493
- 26 Number of Tables: 3
- 27 Number of Figures: 1

28 **Purpose:** Commercially-available microtechnology devices containing 29 accelerometers, gyroscopes, magnetometers, and global positioning technology have 30 been widely used to quantify the demands of rugby union. This study investigated 31 whether data derived from wearable microsensors could be used to develop an 32 algorithm that automatically detects scrum events in rugby union training and match-33 play.

34

Methods: Data were collected from 30 elite rugby players wearing a Catapult S5 Optimeye microtechnology device during a series of competitive matches (n=46) and training sessions (n=51). A total of 97 files were required to "train" an algorithm to automatically detect scrum events using random forest machine learning. A further 310 files from training (n=167) and match-play (n=143) sessions were used to validate the algorithm's performance.

41

42 Results: Across all positions (front row, second row and back row) the algorithm
43 demonstrated good sensitivity (91%) and specificity (91%) for training and match-play
44 events when confidence level of the random forest was set to 50%. Generally, the
45 algorithm had better accuracy for match-play (93.6%) events than training events
46 (87.6%).

47

48 Conclusions: The scrum algorithm was able to accurately detect scrum events for front 49 row, second row and back row positions. However, for optimal results practitioners are 49 advised to use the recommended confidence level for each position to limit false 50 positives. Scrum algorithm detection was better with scrums involving five players or 52 more, and is therefore unlikely to be suitable for scrums involving 3 players (e.g. Rugby 53 Sevens). Additional contact and collision detection algorithms are required to fully 54 quantify rugby union demands.

55

56 Keywords: algorithm; microtechnology; team sport; scrum

57 Commercially-available microtechnology devices containing global positioning 58 systems (GPS) and microsensors (accelerometers, gyroscopes and magnetometers) are 59 commonly used to quantify the physical demands of Rugby Union.¹ During match-play 60 and training, players are divided into subgroups of forwards and backs and are required to perform repeated bouts of high-intensity locomotor activity (sprinting, running, 61 accelerations) separated by low-intensity activity (standing, walking, jogging).¹⁻⁶ In 62 63 addition to the locomotor demands of match-play, players are frequently involved in high-intensity physical contacts and collisions such as mauls, tackles and rucks, with 64 forwards also required to compete in scrums.¹⁻⁸ Scrums are used to restart play after a 65 66 minor infringement and involve all eight forwards from each team, forming three 67 interconnected rows of players. While facing each other, the players forming the front 68 row for each team lock heads and shoulders with the opposition forwards and attempt 69 to produce a greater force than their opponents to gain possession of the ball.⁹

70

71 Despite researchers accurately quantifying the locomotor demands of elite rugby union, 72 contact events such as scrums, rucks, mauls and tackles are usually combined and defined as 'impacts' when using microtechnology.^{1,4,7} Similarly, research evaluating 73 contact events via video-based time-motion analysis has typically categorised these 74 75 incidents as 'high-intensity efforts'³ or 'static exertions'. ^{5,6,8} Success in rugby union frequently depends on the players' ability to tolerate contact events.¹⁰ However, 76 77 research summarising the physical contribution of contact events (scrums, tackles, 78 rucks and mauls) during match-play, either provide a count of the total number of 79 contact events, a rating of the force involved¹, or the total time attributable to 80 collisions.⁸ To date, no research has differentiated between scrums, rucks, mauls and 81 tackles, which inadvertently implies that each form of contact poses an equal physiological stress to the players.¹¹ Classification of each contact would contribute to 82 an improved understanding of the unique stresses associated with each of these collision 83 84 types. In turn, this would potentially assist to improve player preparation and help to 85 reduce the risk of injury and/or re-injury during training and competition.

86

Microsensors have been used to quantify the demands of sport-specific movements in 87 88 team sports, snow sports, individual sports and water sports.¹¹ Validated algorithms 89 have been applied to microsensor data to automate the collection of sport-specific movements, such as cricket fast bowling, ¹² baseball pitching, ¹³ and rugby league 90 tackling. ^{11,14,15} To date, researchers have only used microsensors to quantify the tackle 91 in rugby union,¹⁶ whilst scrums, rucks and mauls have been neglected.¹¹ Researchers 92 have highlighted the injury risk associated with scrums,¹⁷ predominantly in match-93 94 play.¹⁸ Currently there is no other valid method of quantifying scrum workload during training or match-play apart from using video-based time motion analysis, which is a 95 labour-intensive process.¹¹ Many researchers have highlighted the need to further 96 97 investigate contact movements in rugby union, as they generally require the body to 98 endure very high forces that are imparted over a relatively short time period. However, 99 despite the relatively short duration of each contact event, the repeated collisions 100 involved in a typical training or match-play scenario make a significant contribution to 101 the players' total workload. Of the contact movements performed during regular matchplay, scrum events occur around 25 times per game, while depending on playing 102 position, each player will complete approximately 30 rucks and tackles per match.^{5,11,19-} 103 104

105

Given the need for more time-efficient and accurate methods of evaluating the incidence and physical demands of contact events in Rugby Union, this research sought to establish the validity of a microsensor-based algorithm for the automatic detection of scrum events during training and match-play. Based on the demonstrated capabilities of inertial devices to quantify other aspects of sports performance,^{11,22} it was hypothesised that scrum events could be accurately detected using wearable microsensors.

Methods

115 Subjects

116 Thirty elite forwards (mean \pm SD age; 28.3 \pm 4.0 yrs), including players from all 117 positions of the scrum (Front Row, n=16; Second Row, n=8; Back Row, n=6) were 118 recruited to develop and validate the scrum-detection algorithm. At the time of testing, 119 all participants were free of injury and had no known medical conditions that would 120 compromise their participation or influence the recorded outcomes. All participants 121 received a clear explanation of the study's requirements and provided written consent 122 prior to their involvement. The Institution's Human Research Ethics Committee 123 approved all experimental procedures (Approval #2014-135Q).

124

113 114

- 125 Design
- 126 Phase 1 Algorithm Development

127 To facilitate the initial development and training of the scrum detection algorithm, data 128 were collected for the 30 participants using a Catapult S5 Optimeye device (Melbourne, 129 Victoria, Australia) positioned between the players' shoulder blades in a purpose-built 130 vest. Each device contained an array of inertial sensors (i.e. tri-axial accelerometer, 131 gyroscope, magnetometer), which captured data at 100 Hz during a series of 132 competitive matches (n=46) and training sessions (n=51). A total of 97 data files (Front 133 Row, n=49 files; Second Row, n=25 files; Back Row, n=23 files) that captured 1057 134 scrum events were required to develop and optimize the final scrum-detecting 135 algorithm. Timestamps of the scrum instances were manually identified using video 136 data, which were coded alongside Opta Sports events when available (i.e. during match-137 play).

138 The development of an algorithm to detect scrum events involved two separate, but 139 inter-related processes. Firstly, given the unique posture adopted by players while 140 performing scrums, orientation of the device was estimated using a proprietary sensor 141 fusion algorithm that included accelerometer and gyroscope data (Catapult; Melbourne, 142 Victoria, Australia) within a match-play or training session. According to previous research, accelerations and the orientations determined from microsensor data using 143 144 fusion-based methods have excellent reliability and concurrent validity.²³⁻²⁵ While the 145 wearable sensors provided an array of measures, the following criteria were shown to 146 have the ability to identify all scrum instances in the training set and, hence, were the 147 two orientation measures consistently used in the scrum detection algorithm:

148 i. The orientation of the device was below 25 degrees compared to the horizontal
149 plane for at least 4 s. When this criterion was met, the algorithm established this
150 time period as a potential event window.

151 ii. The event was recorded only if the orientation of the device went below the152 horizontal plane during the event window.

153 For data to be considered to potentially represent a true scrum event during training or 154 match-play, both of these orientation criteria were required to be met. This was 155 typically met by participants in preparation for the scrum so that even if a scrum 156 collapsed it would enter the second step of the algorithm. These two initial criteria were 157 intended to remove other non-relevant contact instances. All possible scrum instances 158 within the time-series data were then classified as true and false scrum instances based 159 on video analysis conducted by Opta Sports (http://www.optasports.com) statistics. The window of the classified events were then created for the inertial data and window mid-160 161 points were then extracted to become the event timestamp. This first step of the 162 algorithm development aimed to efficiently transform the data from a time series into 163 a classification problem using the orientation criteria. The second step extracted 164 features of the accelerometer and gyroscope signals from each event. These calculations 165 included summary statistics using different time windows around the event timestamp 166 and formed the variables 33 for the machine-learning process. Variable selection was 167 then performed using the R statistical software package's Variable Selection Using Random Forests (VSURF)²⁶ function. Based on a 10-fold cross-validation mean 168 classification accuracy, 11 signal features were eventually selected from the 169 170 accelerometers and gyroscopes and included in the final version of the random forest classifier.²⁷ R statistical software package (<u>http://www.r-project.org/</u>) was used 171 172 throughout the development of the algorithm.

A scrum confidence scoring was attached to the algorithm based upon the number of trees in the random forest agreeing that a scrum event had taken place. If only the minimum orientation measures were met then the algorithm would return a confidence of 0%. In contrast, when a larger number of trees in the random forest reported detecting a scrum event based on the 11 signal features (Table 1), the algorithm returned a higher confidence rating (maximum 100%).

179 Table 1. List of scrum algorithm signal features

Signal Feature	Feature Name	Feature Description			
1	Horizontal Position 5	To detect how long the estimated orientation of the device is below 5 degrees (i.e. forward flexion)			
2	Horizontal Position 15	To detect how long the estimated orientation of the device is below 15 degrees (i.e. forward flexion)			
3	Horizontal Position 25	To detect how long the estimated orientation of the device is below 25 degrees (i.e. forward flexion), which corresponds with scrum activity			
4	Raw Player Load Q75	75th percentile of raw player load during the scrum activity			

5	Rotation Median	Median of smoothed total rotation during the scrum activity
6	Smooth Player Load 75	75th percentile of smoothed player load during the scrum activity
7	Raw Player Load Q90	90th percentile of raw player load during the scrum activity
8	Raw Player Load Median	Median of raw player load during scrum activity
9	Inertial Side Q10	To detect how long the estimated orientation of the device is below 5 degrees (i.e. forward flexion)
10	Raw Player Load Pre 30	To detect how long the estimated orientation of the device is below 15 degrees (i.e. forward flexion)
11	Raw Rotation Player Load Pre 30	To detect how long the estimated orientation of the device is below 25 degrees (i.e. forward flexion), which corresponds with scrum activity

180

181 Phase 2 – Algorithm Validation

182 To validate the random-forest classifier-based algorithm, a testing set of 21 participants 183 (Front Row, n=9; Second Row, n=5; Back Row, n=7) from the same cohort were monitored using Optimeye S5 devices across 11 international matches (143 full match 184 files) and 9 training sessions (167 full training files). Training session scrums included 185 186 events against opposition (8v8) or against a scrum machine (front 3 against machine, front 5 against machine and 8 against machine). A total of 261 scrum instances 187 (international matches, n=169; training, n=92) were manually coded using video data 188 189 and the timing of each scrum instance was noted according to video, time of day and 190 time on the Catapult raw file. Video coded instances were compared to those detected by the algorithm. Scrum algorithm confidence scoring was set to the lowest possible 191 192 setting, 0%, therefore incorporating all 4833 instances. Each instance was then matched 193 with the relevant time stamp and false positives were thoroughly checked against video 194 coded scrum events.

195

196 Statistical Analysis

197 True positive and negative results and false positive and negative results (Table 2) were 198 determined to calculate algorithm accuracy, precision, specificity and sensitivity. 199 Receiver Operating Characteristic (ROC) analyses were conducted to determine the 200 sensitivity and specificity of the algorithm's confidence in predicting scrum events. The 201 predictive confidence value that yielded the best sensitivity and specificity was selected 202 as the optimal cut-off score and represented the point that simultaneously maximised both on the ROC curve. All statistical analyses were conducted in the StatisticalPackage for the Social Sciences (SPSS v24).

205 **Table 2. Criteria of algorithm results.**

	True	False
Positive	Scrum event and scrum correctly detected	No scrum event, scrum event incorrectly detected
Negative	No scrum event and no scrum event detected	Scrum event and no scrum event detected

206

207 Results

208 To evaluate the performance of the scrum detection algorithm when only the two initial 209 orientation criteria were applied without considering the results of the machine-learning 210 model (i.e. the non-optimised algorithm), the sensitivities and specificities associated with an algorithm confidence of 0% were examined. When data for all positions (i.e. 211 212 front row, second row, back row) and all sessions (i.e. training, competitive matches) 213 were considered, the non-optimised algorithm identified 3904 possible scrum instances. 214 Of these instances, only 25 true negatives were recorded, yielding a sensitivity of 99.5%, a specificity of 31.5% and a precision of 47% (Table 2). Overall, algorithm 215 216 performance was slightly better for match-play (sensitivity 99.8%, specificity 35.0%) 217 than training (sensitivity 98.9%, specificity 28.1%).

Using the 11 signal features identified during the model learning process, the 218 219 algorithm's predictive capacity was improved and this was reflected in the higher 220 predictive confidence values (i.e. the optimised algorithm). Table 3 demonstrates the 221 algorithm confidence cut-offs that returned the best results for the entire dataset and for 222 the three positional groups during the training and match-play sessions based upon 223 receiver operating characteristic analysis (Figure 1). On the basis of these results, the 224 predictive confidence threshold that yielded the best combination of sensitivities and 225 specificities for the entire cohort was 50%, while the optimal cut-off for matches (37%) 226 was somewhat lower than determined for the training data (54%) (Table 2). When the 227 study cohort was subdivided into positional groups, it was shown that the optimal cut-228 off for front row players was 27% for training and 51% for match-play, compared with 229 91% and 49% for the second row. In contrast the predictive confidence values that 230 provided the best sensitivities and specificities for back row players during training and 231 match-play were 63% and 21%, respectively.

Various training scenarios were observed during data collection, involving three, five
and eight players against a scrum machine and opposed "eight verses eight" scrums.
Importantly, the first two scenarios were only included in the validation phase. Scrums
involving the front row only had the lowest sensitivity (50%) and specificity (97%);
this improved when including both the front row and second row (i.e. for five player
scrums), with both positions attaining sensitivity and specificity of 100%. Eight man

scrums against a scrum machine had the highest sensitivity and specificity for all
positions: respective sensitivity and specificity values; front row, 98% and 99%; second
row, 100% and 100%; and back row, 100% and 100%. Opposed scrums in training
involving 16 players (8v8) also demonstrated high sensitivity and specificity for all 3
positions (front row, sensitivity 98% and specificity 99%; second row, sensitivity 100%
and specificity 100%; back row, sensitivity 99.5% and specificity 99.7%).

		Optimal				
	Accuracy (%)	AUC (%)	Cut-Off	Sensitivity	Specificity	
Scrum Identification						
Probability For All Data	91.0	95.8	50%	0.91	0.91	
Data Source						
Probability For Training Data Only	87.6	92.9	54%	0.89	0.87	
Probability For Match Data Only	93.6	98.2	37%	0.94	0.94	
Position						
Probability For Front Row Only	90.4	95 1	41%	0.91	0.90	
Probability For Second Row Only	94.4	97.1	83%	0.94	0.93	
Probability For Back Row Only	89.8	95.8	36%	0.91	0.91	
Position By Data Source						
Probability For Front Row in Training	83.8	88.6	27%	0.84	0.83	
Probability For Second Row in Training	91.4	95.3	91%	0.90	0.90	
Probability For Back Row in Training	90.6	96.1	63%	0.91	0.91	
Probability For Front Row in Matches	95.9	99.1	51%	0.96	0.96	
Probability For Second Row in Matches	98.1	99.7	49%	0.98	0.98	
Probability For Back Row in Matches	89.6	96.6	21%	0.90	0.92	
Position By Data Source (Limited)						
Probability For Front Row in Training	85.2	90.5	39%	0.86	0.86	
Probability For Second Row in Training	91.3	95.3	91%	0.90	0.90	
Probability For Back Row in Training	90.8	96.2	63%	0.91	0.91	

Table 3 – Accuracy, Area Under the Curve (AUC), Optimal algorithm cut-off, sensitivity and specificity for each position during each scenario 244

245 Figure1 –Receiver Operating Characteristic (ROC) analyses for Front Row (A, B),

Second Row (C, D) and Back Row (E, F) players during the training and competitive match scenarios, respectively.



249 **Discussion**

250 This is the first study to investigate the use of microtechnology and associated 251 algorithms to automatically detect scrum events in elite rugby union. Our results 252 demonstrate that scrum events were best detected with high sensitivity and specificity 253 when algorithm confidence level was at 50%, although algorithm performance was 254 better during match-play than training. In training, scrums that involved a minimum of 255 8 players (8 against a machine or contested scrums involving 16 players) returned 256 higher accuracy than those scenarios that involved 3 or 5 players. This finding can be 257 explained by the lack of the latter scenarios in the training phase of the algorithm. 258 Accuracy was best for the front row, with detection of scrum events poorest in the back 259 row. These findings provide a practical and valid method of quantifying scrum events 260 in professional rugby union match-play and training sessions.

261 False negatives during training were only recorded during 3-man scrums performed 262 against a machine. This may have been due to the activity duration being insufficient 263 to satisfy the algorithm's minimum requirements, thus affecting the overall sensitivity 264 and specificity for the front row players during training sessions. Other false negatives 265 in training occurred when scrums collapsed (front row falls to floor) or were reset 266 (incorrect positioning) affecting both the front row and back row. During match-play, 267 all false negatives were attributable to players in the back row who did not maintain a 268 horizontal position for an adequate period of time to satisfy the algorithm's least 269 common denominators before a scrum collapse. As shown in the results for these 270 players, the tendency for back row players to change their trunk orientation prior to a 271 scrum collapse significantly affected the algorithm's sensitivities and specificities for 272 this positional group. Although the results for the back row players were negatively 273 affected by this phenomenon, they do suggest that the physical exertion exhibited by 274 these individuals during a particular scrum event may be quite different to that of front 275 and second row players, even if a scrum is completed or collapses.

276 The comparisons of video-based notational analysis and the scrum algorithm 277 demonstrated the best results with a 50% threshold cut off. The overall outcome of the 278 algorithm was better for match-play than training. Fewer scrum variations occur in 279 match-play (i.e. each scrum is always contested by 16 players), whereas training 280 activities may involve contested '8 v 8' scrums, eight players against a scrum machine, 281 or the front five (involving front row and second row) and front row positions only, 282 which may account for the differences in algorithm performance in different scenarios. 283 Further analysis of the different types of scrum-based technical drills utilized during 284 training indicated that the algorithm performed worse for drills involving only three or 285 five players. Although these results suggest that the algorithm's performance may be 286 improved by including such drills in the "learning" phase of the algorithm, it could be 287 argued that scrums involving 5 or fewer players are aimed more at developing 288 technique, rather than specifically preparing the athletes for the demands of match-play. 289 As such, the specific differences between these training-based drills and actual scrum 290 events may contribute to these incidents not being identified as a scrum using the 291 specified algorithm criteria.

292

We found that algorithm performance differed among positions during match-play and training. Optimal sensitivity and specificity for all positions occurred when the algorithm confidence rating was set at 37% for match-play and 54% for training (Table 296 2). Due to the differences in algorithm performance among positions, setting 297 confidence thresholds of 51%, 49% and 21% during match-play and 27%, 91%, and 298 63% during training for the front row, second row, and back row, will likely produce 299 optimum results, although caution must be taken when extrapolating these results to 300 other independent data sets. False positive events (threshold set to 50%) totaled 168 301 and 1668 true negative events (predominantly scoring below 5% confidence) across the 302 validation data set. Most events were off camera, although events scoring the highest 303 confidence rating were from rare static maul events where players were not moving and 304 positioned in a similar posture to that observed during a scrum

305 The results of the scrum algorithm are in agreement with a recent systematic review 306 that evaluated the use of microsensors for the detection of sport-specific movements.¹¹ This technology has been applied in cricket to count balls bowled¹² and bowling 307 intensity,²⁸ baseball throwing,¹³ tennis serves,²⁹ and several individual,^{11,30-32} snow,^{11,33-38} and water-based^{11,39-41} sports. Microsensors and associated algorithms have been 308 309 used to detect tackles in rugby league¹⁴ with accuracy improving with greater impact 310 forces and longer duration of events.¹⁵ However, this technology has previously been 311 312 shown to be less useful for detecting tackle events in rugby union²¹ and Australian 313 football⁴² match-play. A possible explanation for the poor performance of the algorithm 314 in Australian football and rugby union match-play is that the tackle algorithm was 315 trained on rugby league players, to identify rugby league tackles. The differences in 316 tackles between rugby league and that of Australian football and rugby union may 317 explain the differences in accuracy and show the importance of the representativeness 318 of the training data set for developing movement specific algorithms. Given the 319 differences in findings among rugby league, rugby union, and Australian football, and 320 the present findings that 3- and 5-man scrums were less accurate than 8-man scrums. 321 we would recommend only using the scrum algorithm for detecting scrum events 322 involving 15-a-side rugby union.

323 Although this algorithm advances the ability of sport scientists to automatically detect 324 scrum events in elite rugby union, there are some potential limitations to the research. 325 The algorithm was designed using two elite level teams and tailored primarily for front 326 row players due to their role within scrum events. This may account for the slight, but 327 incremental decrease in algorithm performance for the second row and back row 328 positions, respectively. Elite male players were used to train the algorithm; 329 consequently, the algorithm may be less applicable for younger and smaller junior 330 rugby union participants, or female players, due to possible difference in microsensor 331 signals. Finally, at present, the scrum algorithm only detects the number of scrum 332 events and does not account for the forces applied during these events. Despite these 333 limitations, this study demonstrates the potential for microsensor technology in the 334 detection of rugby union-specific collision events provided an adequate (i.e. specific 335 and representative) training data set. While the demonstrated success of the presented 336 algorithm suggests that practitioners will be better able to detect scrum events in 337 training and match-play to monitor players' total training loads, it is important to 338 acknowledge that the scrum is one of many contact types experienced in rugby union. 339 Hence, despite the algorithm success, a complete understanding of a player's match 340 demands and total training load would require the development of alternate, but 341 complementary methods to identify rucks, tackles and mauls using microtechnology. 342

12

343 PRACTICAL APPLICATIONS

344 The majority of rugby union GPS analyses have focussed on the locomotor demands (i.e. low-speed activities, high-speed running, and sprinting) of the game.¹⁻⁶ However, 345 346 disregarding the physically demanding collision events that may result in very little 347 locomotor activity, may severely underestimate the physical demands of match-play. 348 The development and validation of a scrum algorithm to automatically detect scrum 349 events during training and match-play improves the understanding of an important 350 component of rugby union. Previously, this type of analysis would require time consuming video-based notational analysis. The automated detection of scrum events 351 352 using data provided by the GPS units worn by players allows practitioners to more 353 easily quantify the occurrence of scrum events during regular training and match-play situations. By improving the efficiency of this process, it becomes far more viable for 354 355 sports scientists to determine the physical load associated with these contact events, 356 which should ultimately improve player preparation and reduce the risk of injury. 357 Further research investigating the use of this technology to quantify the ruck, tackle and 358 maul is warranted.

359

360 **CONCLUSION**

361 In conclusion, we investigated the use of microtechnology and associated algorithms to automatically detect scrum events in elite rugby union. Receiver Operating 362 363 Characteristic analyses provided optimal random forest algorithm confidence 364 thresholds to generate best sensitivity and specificity (typically >90%). Algorithm 365 performance was better during match-play than training for front row and second row, 366 although conversely, results revealed better performance for the back row during 367 training than match-play. In training, scrums that involved a minimum of eight players 368 were readily detected, while scrums involving three players were less accurate. Scrums 369 involving five players or more attained markedly better results. Detection was best for 370 the second row, with decreased detection in the front row, with back row positions 371 performing comparatively lower. These findings provide a practical and valid method 372 of quantifying scrum events in professional rugby union match-play and training.

373

374 **ACKNOWLEDGEMENTS**

375 We would like to acknowledge the players used in this research. Special thanks to 376 Catapult and Jerome Durussel for their involvement with their help throughout this

377 project.

378 REFERENCES

- Cunniffe B, Proctor W, Baker J, Davies B. An evaluation of the physiological demands of elite rugby union using global positioning system tracking software. J. *Strength Cond. Res.* 2009;23(4):1195-1203.
- 2. Lacome M, Piscione J, Hager J, Bourdin M. A new approach to quantifying physical
 demand in rugby union. *Journal of Sports Sciences*. 2013;32(3):290-300.
- 384 3. Austin D, Gabbett T, Jenkins D. The physical demands of Super 14 rugby union. J.
 385 Sci. Med. Sport. 2011;14(3):259-263.
- 4. Coughlan G, Green B, Pook P, Toolan E, O'Connor S. Physical game demands in
 elite rugby union: a global positioning system analysis and possible implications for
 rehabilitation. *J Orthop Sports Phys Ther.* 2011;41(8):600-605.
- 389 5. Quarrie K, Hopkins W, Anthony M, Gill N. Positional demands of international
 390 rugby union: Evaluation of player actions and movements. J. Sci. Med. Sport.
 391 2013;16(4):353-359.
- 392 6. Duthie G, Pyne D, Hooper S. Applied physiology and game analysis of rugby
 393 union. *Sports Med.* 2003;33(13):973-991.
- 394 7. Cahill N, Lamb K, Worsfold P, Headey R, Murray S. The movement characteristics
 395 of English Premiership rugby union players. *Journal of Sports Sciences*.
 396 2013;31(3):229-237.
- 8. Roberts S, Trewartha G, Higgitt R, El-Abd J, Stokes K. The physical demands of
 elite English rugby union. *Journal of Sports Sciences*. 2008;26(8):825-833.
- 399 9. Wu W, Chang J, Wu J, Guo L. An investigation of rugby scrimmaging posture and
 individual maximum pushing force. *J. Strength Cond. Res.* 2007;21(1):251-258.
- 401 10. Gabbett T. Quantifying the physical demands of collision sports. *J. Strength Cond.*402 *Res.* 2013;27(8):2319-2322.
- 403 11. Chambers R, Gabbett T, Cole M, Beard A. The use of wearable microsensors to
 404 quantify sport-specific movements. *Sports Med.* 2015;45(7):1065-1081.
- 405 12. McNamara D, Gabbett T, Chapman P, Naughton G, Farhart P. The validity of
 406 microsensors to automatically detect bowling events and counts in cricket fast
 407 bowlers. *Int. J. Sports Physiol. Perform.* 2015;10(1):71-75.
- 408 13. Murray N, Black G, Whiteley R et al. Automatic detection of pitching and throwing
 409 events in baseball with inertial measurement sensors. *Int. J. Sports Physiol.*410 *Perform.* 2016:1-18.
- 411 14. Gabbett T, Jenkins D, Abernethy B. Physical collisions and injury during
 412 professional rugby league skills training. J. Sci. Med. Sport. 2010;13(6):578-583.
- 413 15. Hulin B, Gabbett T, Johnston R, Jenkins D. Wearable microtechnology can
 414 accurately identify collision events during professional rugby league match-play. J.
 415 Sci. Med. Sport. 2017;20(7):638-642.

- 416 16. Kelly D, Coughlan G, Green B, Caulfield B. Automatic detection of collisions in
 417 elite level rugby union using a wearable sensing device. *Sports Engineering*.
 418 2012;15(2):81-92.
- Taylor A, Kemp S, Trewartha G, Stokes K. Scrum injury risk in English
 professional rugby union. *Br. J. Sports Med.* 2014;48(13):1066-1068.
- 421 18. Moore I, Ranson C, Mathema P. Injury risk in international rugby union. *Orthop.*422 *J. Sports Med.* 2015;3(7):232596711559619.
- Hartwig T, Naughton G, Searl J. Motion analyses of adolescent rugby union
 players: a comparison of training and game demands. J. Strength Cond. Res.
 2011;25(4):966-972.
- 426 20. Quarrie K, Raftery M, Blackie J et al. Managing player load in professional rugby
 427 union: a review of current knowledge and practices. *Br. J. Sports Med.*428 2016;51(5):421-427.
- 429 21. Reardon C, Tobin D, Tierney P, Delahunt E. Collision count in rugby union: A
 430 comparison of micro-technology and video analysis methods. *Journal of Sports*431 *Sciences*. 2016:1-7.
- 432 22. Wundersitz D, Josman C, Gupta R, Netto K, Gastin P, Robertson S. Classification
 433 of team sport activities using a single wearable tracking device. *J. Biomech.*434 2015;48(15):3975-3981.
- 435 23. Olivares A, Górriz J, Ramírez J, Olivares G. Using frequency analysis to improve
 436 the precision of human body posture algorithms based on Kalman filters. *Comput.*437 *Biol. Med.* 2016;72:229-238.
- 438 24. Bergamini E, Ligorio G, Summa A, Vannozzi G, Cappozzo A, Sabatini A.
 439 Estimating orientation using magnetic and inertial sensors and different sensor
 440 fusion approaches: accuracy assessment in manual and locomotion tasks. *Sensors*.
 441 2014;14(10):18625-18649.
- 442 25. Cole M, van den Hoorn W, Kavanagh J et al. Concurrent Validity of accelerations
 443 measured using a tri-axial inertial measurement unit while walking on firm,
 444 compliant and uneven surfaces. *PLoS ONE*. 2014;9(5):e98395.
- 445 26. Genuer R, Poggi J, Tuleau-Malot C. VSURF: An R package for variable selection
 446 using random forests. *The R Journal*. 2015;7(2):19-33.
- 447 27. Breiman L. *Machine Learning*. 2001;45(1):5-32.
- 448 28. McNamara D, Gabbett T, Blanch P, Kelly L. The relationship between wearable
 449 microtechnology device variables and cricket fast bowling intensity. *Int. J. Sports*450 *Physiol. Perform.* 2017:1-20.
- 451 29. Whiteside D, Cant O, Connolly M, Reid M. Monitoring hitting load in tennis using
 452 inertial sensors and machine learning. *Int. J. Sports Physiol. Perform.* 2017:1-20.
- 30. Helten T, Brock H, Müller M, Seidel H. Classification of trampoline jumps using
 inertial sensors. *Sports Engineering*. 2011;14(2-4):155-164.

- 455 31. Lai D, Hetchl M, Wei X, Ball K, Mclaughlin P. On the difference in swing arm
 456 kinematics between low handicap golfers and non-golfers using wireless inertial
 457 sensors. *Procedia Eng.* 2011;13:219-225.
- 458 32. Lee J, Mellifont R, Burkett B, James D. Detection of illegal race walking: a tool to
 459 assist coaching and judging. *Sensors*. 2013;13(12):16065-16074.
- 460 33. Beanland E, Main L, Aisbett B, Gastin P, Netto K. Validation of GPS and 461 accelerometer technology in swimming. *J. Sci. Med. Sport.* 2014;17(2):234-238.
- 462 34. Dadashi F, Crettenand F, Millet G, Aminian K. Front-crawl instantaneous velocity
 463 estimation using a wearable inertial measurement unit. *Sensors*.
 464 2012;12(12):12927-12939.
- 35. Dadashi F, Crettenand F, Millet G, Seifert L, Komar J, Aminian K. Automatic frontcrawl temporal phase detection using adaptive filtering of inertial signals. *Journal of Sports Sciences*. 2013;31(11):1251-1260.
- 468 36. Fulton S, Pyne D, Burkett B. Validity and reliability of kick count and rate in
 469 freestyle using inertial sensor technology. *Journal of Sports Sciences*.
 470 2009;27(10):1051-1058.
- 471 37. Fulton S, Pyne D, Burkett B. Quantifying freestyle kick-count and kick-rate patterns
 472 in Paralympic swimming. *Journal of Sports Sciences*. 2009;27(13):1455-1461.
- 38. James D, Leadbetter R, Neeli M, Burkett B, Thiel D, Lee J. An integrated swimming
 monitoring system for the biomechanical analysis of swimming strokes. *Sports Technology*. 2011;4(3-4):141-150.
- 476 39. Chardonnens J, Favre J, Le Callennec B, Cuendet F, Gremion G, Aminian K.
 477 Automatic measurement of key ski jumping phases and temporal events with a
 478 wearable system. *Journal of Sports Sciences*. 2012;30(1):53-61.
- 479 40. Chardonnens J, Favre J, Cuendet F, Gremion G, Aminian K. Characterization of
 480 lower-limbs inter-segment coordination during the take-off extension in ski
 481 jumping. *Hum. Mov. Sci.* 2013;32(4):741-752.
- 482 41. Marsland F, Lyons K, Anson J, Waddington G, Macintosh C, Chapman D.
 483 Identification of cross-country skiing movement patterns using micro-484 sensors. *Sensors*. 2012;12(12):5047-5066.
- 485 42. Gastin P, Mclean O, Breed R, Spittle M. Tackle and impact detection in elite
 486 Australian football using wearable microsensor technology. *Journal of Sports*487 *Sciences*. 2014;32(10):947-953.