Colby, M. J., Dawson, B., Peeling, P., Heasman, J., Rogalski, B., Drew, M. K., Stares, J., Zouhal, H., & Lester, L. (2017). Multivariate modelling of subjective and objective monitoring data improve the detection of non-contact injury risk in elite Australian footballers. *Journal of Science and Medicine in Sport, 20*(12), 1068–1074.

Which has been published in final form at:
https://doi.org/10.1016/j.jsams.2017.05.010

Multivariate modelling of subjective and objective monitoring data improve the detection of non-contact injury risk in elite Australian footballers

M.J. Colby, B. Dawson, P. Peeling, J. Heasman, B. Rogalski, M.K. Drew, J. Stares, Hassane Zouhal, L. Lester

Multivariate modelling of subjective and objective monitoring data improve the detection of non-contact injury risk in elite Australian footballers

Title: Multivariate modelling of subjective and objective monitoring data improve the detection of non-contact injury risk in elite Australian footballers

Authors: Marcus J Colby [a,b], Brian Dawson [a,b], Peter Peeling [a,c], Jarryd Heasman [b], Brent Rogalski [b], Michael K Drew [d,e] ,Jordan Stares [a,b] , Hassane Zouhal [f] and Leanne Lester [a]

Institution and affiliations:

[a] School of Sport Science, Exercise and Health, The University of Western Australia, Australia

[b] West Coast Eagles Football Club, Australia

[c] Western Australian Institute of Sport, Mt Claremont, Australia.

[d] Australian Institute of Sport, Canberra, Australia.

[e] Australian Centre for Research into Injury in Sport and its Prevention (ACRISP), Federation University Australia, Ballarat, Australia.

[f] Movement, Sport and Health Sciences laboratory (M2S). UFR-APS, University of Rennes 2 – ENS Rennes, Avenue Charles Tillon, CS 24414, 35044 Rennes Cedex, France.

Corresponding Author:

Marcus J Colby

Contact details

Tel: +61 448 354 636

Email: marcus_colby@hotmail.com

Abstract

Objectives: To assess the association between workload, subjective wellness, musculoskeletal screening measures and non-contact injury risk in elite Australian footballers.

Design: Prospective cohort study.

Methods: Across 4 seasons in 70 players from one club, cumulative weekly workloads (acute; 1 week, chronic; 2-, 3-, 4-week) and acute:chronic workload ratio's (ACWR: 1-week load/average 4-weekly load) for session-Rating of Perceived Exertion (sRPE) and GPS-derived distance and sprint distance were calculated. Wellness, screening and non-contact injury data were also documented. Univariate and multivariate regression models determined injury incidence rate ratios (IRR) while accounting for interaction/moderating effects. Receiver operating characteristics determined model predictive accuracy (area under curve: AUC).

Results: Very low cumulative chronic (2-,3-,4- week) workloads were associated with the greatest injury risk (univariate IRR=1.71-2.16, 95% CI=1.10-4.52) in the subsequent week. In multivariate analysis, the interaction between a low chronic load and a very high distance (adj-IRR =2.60, 95% CI=1.07-6.34) or low sRPE ACWR (adj-IRR=2.52, 95% CI=1.01-6.29) was associated with increased injury risk. Subjectively reporting *"yes" (vs. "no")* for old lower limb pain and heavy non-football activity in the previous 7 days (multivariate adj-IRR=2.01-2.25, 95% CI=1.02-4.95) and playing experience (> 9 years) (multivariate adj-IRR=2.05, 95% CI=1.03-4.06) was also associated with increased injury risk, but screening data were not. Predictive capacity of multivariate models was significantly better than univariate ($AUC_{multivariate}$=0.70, 95%CI 0.64 to 0.75; $AUC_{univariate}$ range=0.51-0.60).

Conclusions: Chronic load is an important moderating factor in the workload-injury relationship. Low chronic loads coupled with low or very high ACWR are associated with increased injury risk.

Keywords: injury prevention, team sports, load monitoring, acute:chronic workload ratio

Introduction

Sports medicine/science staff must regularly evaluate player injury risk to assess readiness to train and optimise player game availability, as lower injury rates are associated with enhanced team performance (1). In elite Australian football (AF), a typical in-season weekly cycle involves a competitive game, then recovery, training and subsequent matches. Determining player injury risk at commencement of this cycle, by assessing several sport specific risk factors, is critical for weekly planning and (potentially) lower injury incidence. Recently, a revised model of injury aetiology highlighted the inclusion of workload (2), to complement both non-modifiable characteristics (eg, age/playing experience) (3) and modifiable characteristics (eg, strength/flexibility deficits) (4, 5), when identifying multifactorial injury causes (2).

A recent review (6) of workload and injury risk highlighted several independent injury risk factors, including both low (7, 8) and high (9, 10) chronic (multiple weeks) cumulative workloads; high acute:chronic workload ratio's (ACWR) (7, 11); low chronic workloads in combination with high ACWR (11); and large (>1250 AU) week-to-week load changes (9). However, most team sports research has only modelled load independently, not accounting for interaction or moderating effects (12), which may represent a more holistic method for explaining the dynamic and multifactorial nature of injury.

Additional to load monitoring, weekly perceived wellness responses are commonly collected, with irregularities in player profiles warning of potential over-reaching (13). Regular musculoskeletal screening also occurs to determine any significant deviations from baseline scores, to assess the progress of injury rehabilitation programs, and establish future return-to-play status for healthy players (14). While the authors agree with recent commentary on screening tests, this debate is currently limited to periodic physical examinations (PPEs) which form a component of a primary prevention program (14) and ignore the temporal relationship between screening and the date of injury. In the scenario where a player is routinely assessed and compared to deviations (to their norm) the temporal sequence is accounted for in a time-series manner and may offer a solution, particularly when combined with exposure (workload) data. This approach is synonymous to secondary prevention programs which are aimed at detecting subclinical signs and symptoms such that early management can be implemented.

To date, no study has used these combined measures to assess the level of player injury risk in the AF weekly cycle. Therefore, this study aimed to (a) identify the independent injury risk factors collected weekly (subjective wellness, musculoskeletal screening, workload) that predispose an athlete to injury in the subsequent 7 days (15), and (b) establish a multivariate model combining the best injury risk predictors to aid individualised workload management.

Methods

Player data (n=70: 49 players were listed in multiple seasons) from one Australian Football League (AFL) club across four consecutive seasons was used. In total, 3507 individual in-season weekly data points were collected. Mean (± SD) player age, stature and body mass were: 22.9 ± 3.4 y, 188.1 ± 6.6 cm and 87.0 ± 8.2 kg, respectively. For AFL system experience, 23% of players had 1-2 y, 38% had 3-6 y and 39% had 7+ y, respectively. Players either competed in AFL or Western Australian Football League matches across these seasons. All players provided written consent prior to participation. Data was de-identified and extracted from the club's database. Human ethics approval was obtained from the host institution review board (RA/4/1/5015).

Injury information was classified and collated by the club's senior physiotherapist. Here, injury was defined as any lower body non-contact (intrinsic) injury resulting in matches missed (16), since such injuries have been related to training load (17). Non-contact (extrinsic) injuries were not considered.

Training and match workload was defined using both previously validated objective GPS (18) and subjective RPE (19) measures. Multiple external loads were quantified using GPS units (SPI Pro X; GPSports, Canberra, Australia), sampled at an interpolated rate of 15 Hz (true sampling at 5 Hz) and downloaded into a Team AMS analysis program. Distance was defined as total distance covered (m), including walking, running and sprinting. 'Sprint distance' was defined as distance covered (m) above 75% of individual player maximum speed (determined from GPS game data). These commonly used GPS metrics (10,17) were chosen to represent aspects of total and high intensity running volumes within AF demands; other metrics (i.e. additional velocity thresholds, acceleration, deceleration) were not considered due to varying definitions and validation concerns (20).

The "internal" workload was quantified using the "On-Legs sRPE" method, where load (arbitrary units) is the product of the 10-point modified-Borg scale sRPE (9, 19) and total session duration (min). "On-Legs" sessions were defined as any on-field running session where players wore a GPS unit (weights and cross-training data were not available).

Workload data were retrospectively categorised into weekly blocks (Tuesday to following Monday) throughout each season. This structure was chosen as injury risk assessment and subsequent load management would occur following the Monday of each week (club training day without any field training or running). In addition to the weekly acute load (sum of last 7 day period), other load measures were derived using previous studies: a) chronic two, three and four accumulated weekly loads were calculated by summing the previous week's training and game loads (9, 10); b) week-to-week load change (absolute change in current load from previous week) (9); c) ACWR: a player's acute (one week) workload divided by their chronic (four week rolling average) workload (1). Workload category ratings of "very-low" through to "very-high" were created using quintiles, and risk reported in reference to the "moderate load" rating.

Subjective wellness was collected via a customised questionnaire on Mondays; it was brief, specific and based on common components in shortened psychological tools in the literature assessing training imbalances (21). The items included fatigue, sleep quality, muscle soreness, stress levels, mood and perceived performance on a five-point Likert scale, ranging from 1 (as bad as possible) to 5 (as good as possible). Significant wellness declines were calculated as a 1 SD decrease compared to an individual's rolling season-to-date average and SD (13). Further, simple yes/no reporting of questions relating to the past 7 days were considered, including; 'Have you experienced old lower limb pain? (i.e., recurring pain from a previous lower limb injury in the past 12 months); 'Have you completed heavy non-football activities? (i.e., moved house, gardening, painting etc.); and ‘Do you have any lower back pain that is new or worse than last week?’

On Mondays, players also performed several common and validated musculoskeletal screenings (22, 23), including sit and reach (lower back/hamstring flexibility), adductor squeeze (adductor strength) and dorsiflexion lunge left to right differential (ankle stiffness). Although Bahr (14) recently cautioned against using musculoskeletal screening data for injury prediction, our study focused objectively on a significant change (1 SD decrease) in the individual’s current screening results compared to their rolling season average (as per wellness data). Full descriptions of the test procedures, inter-rater and test-retest reliability statistics are presented in supplementary on-line material (Table A).

A mixed model generalized estimating equation (GEE) analysed the relationship between weekly data and injury in the subsequent week, as these analyses can handle panel data (repeated individual measures). This modelling design is supported by cohort studies (7,11,17) and Level 1 evidence (15) showing an association between workload and injury in the subsequent week. For injury risk (injury/no injury in subsequent week), a Poisson log-link regression with robust error estimate, and exchangeable working correlation structure (within the GEE model) was used (24). Incidence rate ratios (IRR) were calculated. Independent (univariate) GEE regression models for each predictor variable were determined, not accounting for other moderating covariates (12). Expanding on previous research (11) investigating subsets of data in workload-injury relationships, an interaction effect between chronic workload and the acute:chronic workload ratio was entered into a multivariate model. 4-week chronic loads were chosen as the best cumulative load predictor (as demonstrated by the highest area under curve: AUC) for inclusion in multivariate models. To simplify models, 4-week chronic load data was dichotomized by the median score (11), to determine a below (low) and above (high) average 4-week chronic workload. A high chronic load and moderate ACWR was defined as the reference group. A final multivariate model then included significant non-workload related predictors from univariate models. Adjusted IRR (adj-IRR) in the multivariate model represent the risk whilst accounting for moderating (12) effects of other variables.

All models were assessed for model fit using in and out of data methods. In-data model detection capacity was assessed by Receiver Operator Characteristics (ROC) curves and compared using “jack-knife method” (25), with Sidak correction to account for multiple comparisons. To evaluate univariate and multivariate model ability to fit out-of-sample data, *k*-fold cross-validation with 10-folds was utilised (26). For comparison, root mean squared error (RMSE) is reported where lower values and less variability between *k*-folds indicate a better fit. All data analysis was performed in Stata 12 (Stata 12 IC, StataCorp, USA). Significance occurred when an IRR 95% CI did not cross 1.00. Injured players’ data for the weeks following injury were excluded until they returned to main (full) training. Extended statistical methods may be found in supplementary online material.

Results

A total of 97 non-contact (intrinsic) lower body injuries were sustained across the four in-season phases (9.8 per 1000 hours) and were subsequently included in the analysis. Descriptive statistics for workload (Table B) and wellness scores (Table C) over the four seasons are presented in supplementary on-line material.

Table 1 presents significant univariate models. A clear association was evident between very low 2-4 week cumulative chronic loads (distance, On-Legs sRPE) and increased injury risk (IRR= 1.54-2.32, 95% CI= 1.10-4.52), compared to moderate loads. A U-shaped relationship was evident with sprint ACWR, indicating increased injury risk for both very low (IRR= 1.83, 95% CI=1.01-3.32) and very high (IRR= 1.90, 95% CI=1.01-3.58) ranges. Player’s reporting “yes” for heavy non-football activity and old lower limb pain were both associated with increased injury risk (IRR= 2.27-2.31; 95% CI=1.11-4.80). Players with > 9 years of playing experience were at twice the risk (IRR= 2.06, 95% CI=1.04-4.22) compared to 1-2 year players. No significant relationship between wellness scores and non-contact injury in the subsequent 7 days was observed. Injury probabilities derived from univariate models displayed poor predictive accuracy (AUC= 0.52 – 0.60).

Insert table 1 about here

Table 2 presents the multivariate model that produced the highest predictive accuracy. The following inferences account for all other variables in this model (adjusted-IRR). A low chronic distance coupled with a very high distance ACWR was associated with increased risk (adj-IRR= 2.60, 95% CI=1.07-6.34) compared to an above average chronic load and moderate ACWR. Conversely, a low On-Legs sRPE chronic load coupled with a low On-Legs sRPE ACWR was associated with increased risk (adj-IRR= 2.52, 95% CI=1.01-6.29) compared to an above average chronic load and moderate ACWR. Other non-workload related variables (playing experience, heavy non-football activity, old lower limb pain) retained significance in the model, presenting similar risks (adj-IRR= 2.02-2.25, 95% CI=1.02-4.95) to their respective univariate models. Figure 1 presents the multivariate predicted injury probability for each variable, whilst accounting for all other variables in the model. Predictive accuracy of the multivariate model (AUC= 0.70, 95%CI 0.64 to 0.75) was significantly (x^2 = 37.90; p< 0.001) better than all univariate models (AUC= 0.52-0.60) when tested on in-sample data. However, cross fold validation results indicated a very similar fit (k= 10: $RMSE_{univariate}$ mean±SD = 0.16±0.02 compared to $RMSE_{multivariate}$ mean±SD= 0.16±0.02) on out-of-sample data, demonstrating an equal (clinical) ability to predict injury in the subsequent week.

Insert table 2 and figure 1 about here

Discussion

We believe this is the first study to identify a multifactorial (workload, subjective wellness, and player experience) injury risk model in elite Australian footballers. This paper further supports the view that injuries are produced from a complex “web of determinants” (27), with potential moderating (12) effects occurring between these determinants. The theory that 'training load errors' (1, 28) may cause injury due to players being ill-prepared for the demands of the week is supported. However, since most in-season weekly load is derived from games, load errors here (e.g. excessive spikes in game loads compared to recent matches) may also be a key contributor to injury. As previously reported (11), a clear relationship between independently modelled very low cumulative chronic loads and increased injury risk in the subsequent week was also identified. However, both screening (14) and wellness 'red flags' (13) did not improve injury model predictive accuracy here, despite applying an objective, individualized criterion (a 1 SD decline from the norm). Manipulating training loads in response to wellness (13) and/or screening profiles is common in elite sport, possibly explaining the null predictive value, presenting a potential limitation to research designs in these settings.

Several factors may also interact with the workload-injury relationship (12) and may act as mediators or moderators of risk when considered in combination. Through a multivariate approach, inferences can be made whilst accounting for other workload (internal and external) variables and time invariant factors (playing experience). As with Williams et al. (29), 4-week cumulative chronic loads showed the greatest association with injury and were selected to further explore the interaction between chronic load and the ACWR (11) in a multivariate model.

A low chronic load coupled with a very high ACWR (sprint distance) was associated with the greatest injury risk in the subsequent week. Sudden load increases have previously been associated with increased injury risk in the following week (7, 11), with a high chronic (i.e. high 'fitness') distance providing protection for moderate-high ACWR, but increased risk for very high ACWR in elite rugby league players (11). Conversely, a high chronic load coupled with a very high ACWR was not associated with increased risk in this AF cohort. Potentially, players with a high chronic base had less likelihood of an elevated ACWR, since a much greater acute load is necessary to elicit a similar ACWR to those with a low chronic base. Additionally, in elite settings when players with high chronic load experience acute spikes, load management strategies may be implemented to mitigate the 7 day injury risk latent period investigated here. Interestingly, a low OnLegs sRPE chronic load coupled with a very low or low ACWR was associated with 1.6-2.5 times greater risk, compared to a high chronic load and moderate ACWR. Possibly, players who experienced substantial de-loading may have further reduced their chronic load foundation, a scenario shown here to elevate risk, or were susceptible to large acute increases in load (i.e., a sessional spike) during training or game sessions within the 7 day injury lag period investigated here. These findings support previous reports (11) that high chronic loads provide protection when exposed to a very high ACWR.

Interestingly, players reporting old lower limb pain and heavy non-football activity were associated with twice the injury risk, highlighting the contribution of subjective measures in elite environments. While not addressed in this paper, it is hypothesized that these findings are indicating that a recent history of pain may precede an injury incident or may represent a situation where an athlete is hyperalgesic in their response peripheral stimuli. Further, players with > 9 years of playing experience were associated with a greater injury risk, emphasising the importance of managing older players. These variables were also retained in the multivariate model, thereby warranting further investigation to determine the mediating step (12) that may explain the injury association.

Another novel aspect of this study was comparing univariate and multivariate model predictive accuracy on in-sample data. As suggested previously (27), injury may be attributed to a complex "web of determinants", therefore it is unsurprising to find greater accuracy for the multivariate model. However, when tested on out-of-sample data (through cross validation), similar model fit errors were observed between univariate and multivariate approaches, highlighting the challenge of applying these models to derive out-of-sample injury risk.

Several limitations should be acknowledged in this study. Although the predictive accuracy of the multivariate model may be deemed sufficient (AUC= 0.70) (30), predicted probabilities were tested on the fully trained data set (in-sample testing). Greater external validity may be gained by testing the models identified here on larger out-of-sample datasets (i.e., from other AFL teams), or for full prospective seasons. Pre-season training phase data was not analysed due to difficulties in calculating load measures with retrospective calculations (cumulative workloads, ACWR) around the off-season and Christmas breaks. Further, only on-legs field and game loads were calculated; further research may examine methods to quantify total load (i.e., including resistance and cross training) and load foundation achieved during pre-season (17). Lastly, injury was modelled across a 7 day latent period; future analysis may model on a sessional basis to ensure all workload data is captured prior to injury.

Conclusions

Modelling combined injury risk factors is important for assessing the interaction and moderating nature of multiple risk factors. In the models presented here, a player's chronic load greatly influenced the ACWR-injury relationship. A low chronic ("fitness") load coupled with a large acute de-load (< ~0.80 of a 4-week chronic load) or spike (>~1.20-1.40 of a 4-week chronic load) should be considered as potential injury risk factors. For these high risk scenarios, players may further decrease their chronic load foundation, resulting in an underprepared state for competitive demands or their "fatigue" (acute load) outweighs their "fitness"; leading to (overload) injury. Furthermore, simple "yes/no" wellness responses may have predictive value and should be factored into weekly injury risk assessment in elite sport. The findings here can encourage practitioners to embrace the complexity of injury prediction and consider using a multifactorial approach.

Practical Implications

- Multivariate injury risk modelling may increase predictive accuracy by considering the interaction and moderating effects of common risk factors.
- A player's chronic workload foundation plays a large moderating role when modelling injury risk in elite Australian footballers.
- Low acute:chronic workload ratios should be considered an injury risk factor for the subsequent week, as this may lower the chronic load foundation.
- Simple yes/no subjective wellness responses may have injury predictive value in the subsequent week.

Acknowledgements

The Australian Collaboration for Research into Injury in Sport and its Prevention (ACRISP) is one of the International Research Centres for Prevention of Injury and Protection of Athlete Health supported by the International Olympic Committee (IOC).

REFERENCES

1. Gabbett, TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med* 2016. Published Online First: [12/01/16] doi:10.1136/bjsports2015-095788.

2. Windt J, Gabbett TJ. How do training and competition workloads relate to injury? The workload-injury aetiology model. *Br J Sports Med* 2016. Published Online First: [14/07/16] doi:10.1136/bjsports- 2016-096040.

3. Fortington LV, Berry J, Buttifant D, et al. Shorter time to first injury in first year professional football players: A cross-club comparison in the Australian Football League. *J Sci Med Sport* 2016; 19(1): 18-23.

4. Bradley, PS, Portas, MD. The relationship between pre-season range of motion and muscle strain injury in elite soccer players. *J Strength Cond Res* 2007; 21 (4), 1155-1159.

5. Henderson, G, Barnes, CA, Portas, MD. Factors associated with increased propensity for hamstring injury in English Premier League soccer players. *J Sci Med Sport* 2010; 13: 397–402.

6. Drew MK, Cook J, Finch CF. Sports-related workload and injury risk: simply knowing the risks will not prevent injuries. *Br J Sports Med* 2016. Published Online First: [10/05/16] doi:10.1136/bjsports- 2015-095871.

7. Hulin BT, Gabbett TJ, Blanch P, et al. Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *Br J Sports Med* 2014; 48: 708-712.

8. Cross MJ, Williams S, Trewartha G, et al. The influence of in-season training loads on injury risk in professional rugby union. *Int J Sports Physiol Perform* 2016; 11: 350–355.

9. Rogalski B, Dawson B, Heasman J, et al. Training and game loads and injury risk in elite Australian footballers. *J Sci Med Sport* 2013; 16: 499–503.

10. Colby MJ, Dawson B, Heasman J, et al. Accelerometer and GPS-derived running loads and injury risk in elite Australian footballers. *J Strength Cond Res* 2014; 28: 2244–2252.

11. Hulin BT, Gabbett TJ, Lawson DW. The acute:chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *Br J Sports Med* 2016; 50: 231–236.

12. Windt J, Zumbo BD, Sporer B, et al. Why do workload spikes cause injuries, and which athletes are at higher risk? Mediators and moderators in workload–injury investigations. *Br J Sports Med* 2017. Published Online First [08/04/17] doi: http://dx.doi.org/10.1136/bjsports-2016-097255

13. Gallo TF, Cormack SJ, Gabbett TJ, et al. Self-reported wellness profiles during the competition phase of the season. *J Strength Cond Res* 2016. Published Online First [25/05/16] doi: 10.1519/JSC.0000000000001515.

14. Bahr R. Why screening tests to predict injury do not work – and probably never will...: a critical review. *Br J Sports Med* 2016. Published Online First: [25/04/16] doi:10.1136/bjsports- 2016-096256.

15. Drew MK, Finch CF. The relationship between training load and injury, illness and soreness: A systematic and literature review. *Sports Med* 2016; 46(6): 861-83.

16. Orchard J, Hoskins W, Chiro M. Debate: Consensus injury definitions in team sport should focus on missed playing time. *Clin J Sport Med* 2007; 17:192-196.

17. Windt J, Gabbett TJ, Ferris D, et al. Training load-injury paradox: is greater preseason participation associated with lower in-season injury risk in elite rugby league players?. *Br J Sports Med* 2016. Published Online First [13/04/16] doi: 10.1136/bjsports-2016- 095973.

18. Waldron, M, Worsfold, P, Twist, C, et al. Concurrent validity and test-retest reliability of a global positioning system (GPS) and timing gates to assess sprint performance variables. *J Sport Sci* 2011; 29: 1613–1619.

19. Impellizzeri FM, Rampinini E, Coutts AJ, et al. Use of RPE-based training load in soccer. *Med Sci Sports Exerc* 2004; 36(6): 1042-7.

20. Malone JJ, Lovell R, Varely MC, et al. Unpacking the black box: Applications and considerations for using GPS devices in sport**.** *Int J Sports Physiol Perform* 2016. Published Online First [02/10/16] doi: http://dx.doi.org/10.1123/ijspp.2016-0236.

21. Gastin PB, Meyer D, Robinson D. Perceptions of wellness to monitor adaptive responses to training and competition in elite Australian football. *J Strength Cond Res* 2013; 27(9): 2518-2526.

22. Dennis RJ, Finch CF, Elliott BC, et al. The reliability of musculoskeletal screening tests used in cricket. *Phys Ther Sport* 2008; 9(1): 25-33

23. Gabbe BJ, Bennell KL, Wajswelner H, et al. Reliability of common lower extremity musculoskeletal screening tests. *Phys Ther Sport* 2004; 5(2): 90-97

24. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;159(7):702-6.

25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988:837-845.

26. Schonlau M. Boosted regression (boosting): An introductory tutorial and a Stata plugin. *Stata Journal* 2005; 5(3):330.

27. Bittencourt NFN, Meeuwisse WH, Mendonça LD, et al. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. *Br J Sports Med* 2016; 50: 1309– 1314.

28. Drew MK, Purdam C. Time to bin the term 'overuse" injury/ is 'training load error' a more accurate term? *Br J Sports Med* 2016. Published Online First: [03/02/16] doi:10.1136/bjsports-2015-095543.

29. Williams S, Trewartha G, Cross MJ, et al. Monitoring what matters: A systematic process for selecting training load measures. *Int J Sports Physiol Perform* 2016. Published Online First: [04/10/16] doi: http://dx.doi.org/10.1123/ijspp.2016-0337.

30. Okeh UM, Okoro, CN. Evaluating measures of indicators of diagnostic test performance: Fundamental meanings and formulas. *J Biom Biostat* 2012; 3: 132.

Table 1. Univariate Models: Injury likelihood in subsequent week

	Unit	% INJ	IRR (95% CI)	AUC	RMSE (mean ± SD)
2-week Distance					
Very Low	< 34927 m	4.2	1.54 (0.83-2.84)		
Low	34927 - 39666 m	2.4	0.88 (0.46-1.69)		
Moderate (reference)	39666 - 43179 m	2.7	1.00	*0.59*	0.16 ± 0.03
High	43179 - 47220 m	2.8	1.06 (0.58-1.94)		
Very High	> 47720 m	1.5	0.59 (0.29-1.22)		
3-week Distance					
Very Low	**< 52947 m**	**4.6**	**2.15 (1.15-4.01)**		
Low	52947 - 59077 m	2.8	1.31 (0.75-2.29)		
Moderate (reference)	59077 - 64053 m	2.1	1.00	*0.60*	0.16 ± 0.02
High	64053 - 69042 m	2.2	1.06 (0.50-2.26)		
Very High	> 69042 m	1.7	0.80 (0.37-1.72)		

4-week Distance					
Very Low	**< 71059 m**	**4.6**	**2.32 (1.19-4.52)**		
Low	71059 - 78627 m	3.1	1.56 (0.81-3.03)		
Moderate (reference)	78627 - 84879 m	1.9	1.00	*0.60*	0.16 ± 0.02
High	84879 - 91013 m	1.9	1.00 (0.46-2.21)		
Very High	> 91013 m	1.8	0.94 (0.42-2.11)		
Distance ACWR					
Very Low	< 0.88	2.8	1.17 (0.63-2.19)		
Low	0.88 -0. 99	1.4	0.60 (0.28-1.32)		
Moderate (reference)	0.99 – 1.08	2.4	1.00	*0.58*	0.16 ± 0.02
High	1.08 – 1.21	3.2	1.37 (0.72-2.59)		
Very High	> 1.21	3.6	1.53 (0.84-2.76)		
2-week Sprint					
Very Low	< 314 m	3.6	1.14 (0.61-2.12)		
Low	314 - 478 m	2.8	0.87 (0.78-1.71)		
Moderate (reference)	478 – 641 m	3.2	1.00	*0.58*	0.16 ± 0.02
High	641 - 832 m	2.4	0.73 (0.36-1.49)		
Very High	**> 832 m**	**1.5**	**0.48 (0.24-0.97)**		
3-week Sprint					
Very Low	< 494 m	3.9	1.77 (0.99-3.15)		
Low	494 - 720 m	2.8	1.26 (0.62-2.58)		
Moderate (reference)	720 – 942 m	2.2	1.00	*0.57*	0.16 ± 0.03
High	942 - 1215 m	2.6	1.18 (0.63-2.22)		
Very High	> 1215 m	1.8	0.82 (0.43-1.60)		
4-week Sprint					
Very Low	< 683 m	3.2	0.86 (0.47-1.56)		
Low	683 - 968 m	2.5	0.67 (0.38-1.17)		
Moderate (reference)	968 – 1247 m	3.8	1.00	*0.58*	0.16 ± 0.02
High	1247 - 1583 m	2.2	0.59 (0.32-1.07)		
Very High	**> 1583 m**	**1.7**	**0.45 (0.25-0.84)**		
Sprint distance ACWR					
Very Low	< 0.67	3.6	1.83 (1.01-3.32)		
Low	0.67 - 0.93	1.9	0.99 (0.50-1.94)		
Moderate (reference)	0.93 - 1.13	2.0	1.00	*0.58*	0.16 ± 0.04
High	1.13 – 1.40	2.1	1.06 (0.55-2.07)		
Very High	**> 1.40**	**3.8**	**1.90 (1.01-3.58)**		
1-week On-Legs RPE					
Very Low	< 775 AU	3.9	1.64 (0.99-2.71)		
Low	775 - 1232 AU	2.8	1.16 (0.60-2.26)		
Moderate (reference)	1232 - 1376 AU	2.4	1.00	*0.57*	0.16 ± 0.03
High	1376 - 1503 AU	2.8	1.18 (0.67-2.08)		
Very High	> 1503 AU	1.7	0.71 (0.34-1.48)		
2-week On-Legs RPE					
Very Low	**< 1760 AU**	**4.7**	**2.00 (1.15-3.46)**		
Low	1760 - 2220 AU	1.8	0.77 (0.38-1.56)		
Moderate (reference)	2220 - 2608 AU	2.4	1.00	*0.60*	0.16 ± 0.02
High	2608 - 2885 AU	2.5	1.09 (0.60-2.08)		
Very High	> 2885 AU	2.1	0.90 (0.50-1.61)		
3-week On-Legs RPE					
Very Low	**< 2752 AU**	**4.0**	**1.69 (1.10-2.62)**		
Low	2752 – 3298 AU	2.6	1.11 (0.57-2.14)		
Moderate (reference)	3298 - 3746 AU	2.4	1.00	*0.59*	0.16 ± 0.02
High	3746 - 4197 AU	2.9	1.25 (0.72-2.15)		
Very High	> 4197 AU	1.4	0.59 (0.31-1.14)		
4-week On-Legs RPE					
Very Low	**< 3688 AU**	**4.5**	**1.59 (1.11-2.66)**		
Low	3688 - 4410 AU	3.5	1.00 (0.61-1.63)		
Moderate (reference)	4410 - 4908 AU	2.6	1.00	*0.61*	0.16 ± 0.02
High	4908 - 5446 AU	1.9	0.56 (0.27-1.13)		
Very High	> 5446 AU	1.4	0.70 (0.33-1.31)		
On-Legs RPE ACWR					
Very Low	< 0.86	4.1	1.38 (0.83-2.30)		
Low	0.86 – 1.02	2.4	1.02 (0.57-1.83)		
Moderate (reference)	1.02 - 1.14	2.3	1.00	*0.53*	0.16 ± 0.02
High	1.14 - 1.30	3.2	1.01 (0.53-1.92)		
Very High	> 1.30	2.1	0.93 (0.48-1.80)		
Playing Experience					

	1 -2 y	2.1	1.00	*0.56*	0.16 ± 0.02
	3 - 4	2.5	1.17 (0.64-2.15)		
	5 - 6	2.6	1.22 (0.59-2.50)		
	7 - 9	2.5	1.20 (0.67-2.14)		
	> 9 y	**4.3**	**2.06 (1.04-4.22)**		
Heavy Non-Football Activity					
	No (reference)	2.6	1.00	*0.52*	0.16 ± 0.01
	Yes	**5.8**	**2.31 (1.11-4.80)**		
Old Lower Limb Pain					
	No (reference)	2.4	1.00	*0.55*	0.16 ± 0.02
	Yes	**5.5**	**2.27 (1.34-3.86)**		

INJ= injured; IRR = incidence risk ratio; CI= confidence interval; SD = standard deviation m= meters; AU = arbitrary unit;

ACWR = acute:chronic workload ratio

Note, predictors where the 95% CI did not cross 1.00 appear in bold.

Table 2. Multivariate Model: Injury likelihood in subsequent week

	Unit		% INJ	adj-IRR (95% CI)	Model AUC	Model RMSE (mean ± SD)
Sprint distance chronic load # acute:chronic workload ratio	*4-week cumulative chronic load*	*ACWR*			0.70	0.16 ± 0.02
Low /Very Low		< 0.67	3.8	1.44 (0.64-3.27)		
Low /Low		0.67 - 0.93	2.0	0.87 (0.31-2.43)		
Low /Moderate	< 1097 m	0.93 - 1.13	1.8	0.78 (0.28-2.16)		
Low /High		1.13 - 1.40	2.7	1.23 (0.51-2.93)		
Low /Very High		> 1.40	4.4	1.60 (0.68-3.78)		
High /Very Low		< 0.67	3.4	1.64 (0.65-4.11)		
High /Low		0.67 - 0.93	1.9	1.00 (0.42-2.41)		
High /Moderate (reference)	> 1097 m	0.93 - 1.13	2.0	1.00		
High /High		1.13 - 1.40	1.6	0.73 (0.29-1.83)		
High /Very High		> 1.40	2.6	0.91 (0.36-2.29)		
Distance chronic load # acute:chronic workload ratio						
Low /Very Low		< 0.88	3.5	1.11 (0.41-2.98)		
Low /Low		0.88 - 0.99	1.8	0.80 (0.20-3.26)		
Low /Moderate	< 81694 m	0.99 - 1.08	3.5	1.62 (0.53-4.89)		
Low /High		1.08 - 1.21	3.5	1.73 (0.72-4.11)		
Low /Very High		**> 1.21**	**4.4**	**2.60 (1.07-6.34)**		
High /Very Low		< 0.88	2.0	0.89 (0.29-2.74)		
High /Low		0.88 - 0.99	1.1	0.68 (0.19-2.42)		
High /Moderate (reference)	> 81694 m	0.99 - 1.08	1.5	1.00		
High /High		1.08 - 1.21	3.0	2.16 (0.78-6.02)		
High /Very High		> 1.21	1.6	1.36 (0.32-5.78)		
OnLegs sRPE chronic load # acute:chronic workload ratio						
Low /Very Low		< 0.86	4.2	1.62 (0.70-3.77)		
Low /Low		**0.86 – 1.02**	**6.9**	**2.52 (1.01-6.29)**		
Low /Moderate	< 4660 AU	1.02 - 1.14	3.4	1.30 (0.37-4.63)		
Low /High		1.14 - 1.30	3.2	1.02 (0.44-2.34)		
Low /Very High		> 1.30	2.5	0.61 (0.26-1.45)		
High /Very Low		< 0.86	1.8	0.86 (0.30-2.50)		
High /Low		0.86 – 1.02	1.5	0.83 (0.34-2.04)		
High /Moderate (reference)	> 4660 AU	1.02 - 1.14	2.2	1.00		
High /High		1.14 - 1.30	1.8	0.67 (0.25-1.85)		
High /Very High		> 1.30	1.9	0.62 (0.13-3.07)		
Playing Experience						
	1 – 2 y (reference)		2.1	1.00		
	3 – 4		2.5	1.39 (0.73-2.63)		
	5 – 6		2.6	1.28 (0.59-2.75)		
	7 – 9		2.5	1.37 (0.77-2.43)		
	> 9 y		**4.3**	**2.05 (1.03-4.06)**		
Heavy Non-Football						
	No (reference)		1.0	1.00		
	Yes		**2.3**	**2.02 (1.17-3.49)**		
Old Lower Limb Pain						
	No (reference)		1.8	1.00		
	Yes		**8.8**	**2.25 (1.02-4.95)**		

% INJ = percentage injured; adj-IRR = adjusted incidence rate ratio; CI = confidence interval; AUC = area under curve; RMSE = root mean squared error; m = metres; AU = arbitrary unit; y = years; # = interaction

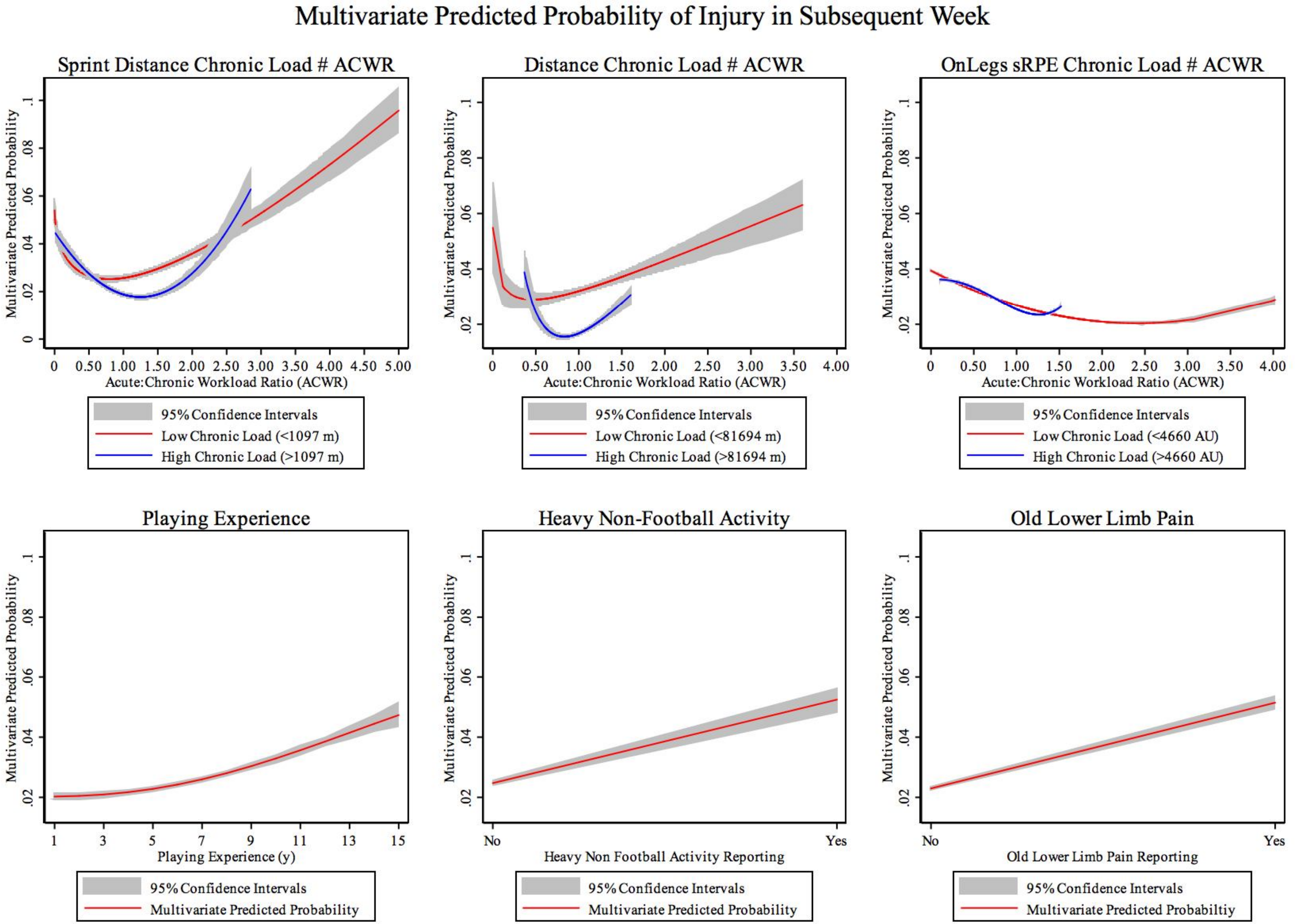
Multivariate Predicted Probability of Injury in Subsequent Week
Sprint Distance Chronic Load # ACWR
Distance Chronic Load # ACWR
OnLegs sRPE Chronic Load # ACWR
Multivariate Predicted Probability
0
.02
.04
.06
.08
.1
0 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4.00 4.50 5.00
0 0.50 1.00 1.50 2.00 2.50 3.00 3.50 4.00
Acute:Chronic Workload Ratio (ACWR)
95% Confidence Intervals
Low Chronic Load (<1097 m)
High Chronic Load (>1097 m)
Low Chronic Load (<81694 m)
High Chronic Load (>81694 m)
Low Chronic Load (<4660 AU)
High Chronic Load (>4660 AU)
Playing Experience
Heavy Non-Football Activity
Old Lower Limb Pain
1 3 5 7 9 11 13 15
Playing Experience (y)
No
Yes
Heavy Non Football Activity Reporting
Old Lower Limb Pain Reporting
Multivariate Predicted Probability
Multivariate Predicted Probabiltiy

Figure Legend

Figure 1. Predicted injury probabilities from multivariate model for all variables (Table 2). The model predicts the probability a player will sustain a non-contact injury in the subsequent week, accounting for interaction (chronic load and ACWR) and moderating (heavy non-football activity, old lower limb pain, playing experience) effects of other variables.