http://dx.doi.org/10.1123/ijspp.2015-0330

Thornton, H.R., Delaney, J.A., Duthie, G.M., Scott, B.R., Chivers, W.J., Sanctuary, C.E. and Dascombe, B.J. (2016) Predicting Self-Reported Illness for Team-Sport Athletes. International Journal of Sports Physiology and Performance 11 (4). pp. 543-550.

Title: Predicting self-reported illness for professional team-sport athletes

Submission type: Original investigation

Authors:
Heidi R. Thornton[1,2], Jace A. Delaney[1,2], Grant M. Duthie[2], Brendan R. Scott[1], William J. Chivers[1], Colin E. Sanctuary[2] and Ben J. Dascombe[1,3]

Institutions:
[1] Applied Sports Science and Exercise Testing Laboratory, Faculty of Science and Information Technology, University of Newcastle, Ourimbah, NSW 2258
[2] Newcastle Knights Rugby League Club, Mayfield, NSW 2304
[3] Priority Research Centre in Physical Activity and Nutrition, University of Newcastle, Callaghan, NSW 2258

Corresponding Author:
Heidi R. Thornton
School of Environmental and Life Sciences,
Faculty of Science and Information Technology,
University of Newcastle
PO Box 127, Ourimbah, NSW 2258
Ph: +61 2 43484149
Email: heidi.thornton@uon.edu.au

Running head: Contributing factors to illness

Abstract word count: 228
Text-only word count: 3303
Number of tables: 2
Number of figures: 4

Abstract

Purpose: The aim of this study was to identify contributing factors to the incidence of illness for professional team-sport athletes, utilizing training load (TL), self-reported illness and wellbeing data. **Methods:** Thirty-two professional rugby league players (26.0 ± 4.8 yr; 99.1 ± 9.6 kg; 1.84 ± 0.06 m) were recruited from the same club. Players participated in prescribed training and responded to a series of questionnaires to determine the presence of self-reported illness and markers of wellbeing. Internal-TL was determined using the session rating of perceived exertion (sRPE) method. These data were collected over 29 weeks, across the preparatory and competition macrocycles. **Results:** The predictive models developed recognized increases in internal-TL (strain values of >2282 AU, weekly-TL >2786 AU and monotony >0.78 AU) to best predict when athletes are at increased risk of self-reported illness. In addition, a reduction in overall wellbeing (<7.25 AU) in the presence of increased internal-TL as previously stated, was highlighted as a contributor to self-reported illness occurrence. **Conclusions:** These results indicate that self-report data can be successfully utilized to provide a novel understanding of the interactions between competition-associated stressors experienced by professional team-sport athletes and their susceptibility to illness. This may assist coaching staff to more effectively monitor players during the season and to potentially implement preventative measures to reduce the likelihood of illnesses occurring.

Keywords: URTI; predictive modelling; rugby league; wellness; sRPE

Introduction

In order to maximize performance and limit any detrimental effects associated with excessive exercise stress in team-sport athletes, it is pivotal for coaching staff to understand the optimal dose-response relationship between exercise training and adaptation.[1,2] It is well established that frequent high-intensity and high-volume training can result in physiological disturbances in an athlete.[3,4] These responses cause short-term reductions in performance, therefore adequate recovery is necessary to elicit the fundamental supercompensation response prior to subsequent training bouts.[1,5] Importantly, if a period of intense training is not well tolerated by an athlete, the incidence of illness and injury is likely to be increased.[6,7]

More specifically, the occurrence of illnesses is common in high-level athletes due to the well-established association between large increases in training load (TL) and suppressed immune function.[8-10] This can be exacerbated by competition demands such as psychological stress, reduced sleep and sub-optimal nutritional intake, which may further inhibit an individual's immune functioning.[11] The presence of illnesses such as upper respiratory tract infections (URTIs) is of great concern for team-sport athletes, given their infectious nature and the potential negative influence on performance and wellbeing.[12] Thus, these athletes should be appropriately monitored to identify periods when they are at increased risk of illnesses.[10,13,14]

There is limited research describing the influence of factors such as training and competition demands on the incidence of illness in professional team-sport athletes. Moreover, the research available on this topic has predominantly involved collecting biochemical markers of endocrine and immune status, which is logistically impractical in high performance team-sport environments.[16,17] The development of more practical and non-invasive methods to assess an athlete's fatigue-recovery profile in this context is therefore of great interest.

Subjective wellbeing questionnaires are recognized for their ability to assess fatigue and psychological variables (e.g. mood and attitude), and provide information regarding individual responses to exercise stress.[2,18] While these questionnaires are now widely used by strength and conditioning coaches to assess how athletes are coping with the stress of training,[19] research has not yet assessed their ability to predict when athletes are at increased risk of illness susceptibility.

Therefore, the purpose of this investigation was to ascertain whether self-report athlete monitoring tools are able to predict periods of increased risk of an illness. It was hypothesized that reductions in perceived wellbeing, in concert with heightened internal-TL scores and periods of stressful competition, would be probable predictors for the incidence of illness in professional team-sport athletes.

Methods

Subjects

Thirty-two professional rugby league players (age = 26.0 ± 4.8 years; body mass = 99.1 ± 9.6 kg and height = 1.84 ± 0.06 m) were recruited from a club competing in the National Rugby League (NRL) competition (n = 18) or the New South Wales (NSW) Cup second-tier competition (n = 14). Participants comprised of a range of playing positions, including forwards (n = 13), backs (n = 10) and adjustables (n = 9) who performed the same training together as a squad.

Design

A longitudinal research design was conducted, whereby self-reported illness, internal-TL for training sessions and match-play, and perceptual wellbeing ratings were collected for 29 weeks during the 2014 Australian Rugby League season. This period included the preparatory and competitive macrocycles. These data were used to develop predictive models to identify factors contributing to the incidence of self-reported illness. Prior to the commencement of the study, all players were provided with information detailing the aims and requirements of the study and provided informed consent. The Institutional Human Ethics Committee approved the research and experimental procedures.

Methodology

Training Requirements

A periodized game-specific training program was prescribed and completed at the discretion of coaching staff. The program varied according to the specific objectives for each macrocycle and the scheduling of competitive matches. The preparatory period (November to February) aimed to apply a demanding training stimulus to develop physiological capacities required for match-play. The training program progressed from high-volume and low-intensity exercise during the preparatory period, to lower-volume and higher-intensity training during the competitive period (March to October). During the competition period the focus of training was to maintain capacities developed during preparation, whilst incorporating game-specific skills and post-match recovery. Table 1 provides a brief overview of the yearly training plan prescribed for professional rugby league players.

INSERT TABLE 1 NEAR HERE

Quantification of Training Loads

The intensity of individual training sessions was estimated using the Category Ratio-10 rating of perceived exertion (RPE) scale.[20] This scale requires athletes to rate the global intensity of the entire training session on a scale from 0-10, where a score of 0 denotes complete rest and 10 indicates maximal effort.[20] Athletes provided RPE scores individually to a member of the research team at 30 minutes following each training session to eliminate any affect that the final phase of training may have on scores.[21] Internal-TL was subsequently calculated by multiplying the RPE value for each player by the duration of the training session (minutes), to provide an index of TL in arbitrary units. Training monotony and strain values were also calculated by dividing the weekly internal-TL by the standard deviation of the individual's weekly-TL (monotony), and then multiplying the athlete's weekly-TL and monotony values (strain).[22]

Self-Reported Information

A multi-component wellness questionnaire (Figure 1) was implemented as part of the club's monitoring practices, adapted from those used in previous research.[18,23] The two-component questionnaire was completed by athletes on a weekly basis, prior to the first training day of the week. It was completed at the same time of day in order to minimize the potential for diurnal fluctuations in mood state and perceived wellness. The self-reported presence of an illness (Part A) was recorded, whereby athletes' documented whether they were suffering from pre-determined symptoms common to that of URTI.[10] Part B evaluated the athlete's perceived wellness and severity of muscle soreness using a 10-point Likert scale ranging from poor to excellent. Participants were asked to firstly rate their nutritional intake, based on how closely they had adhered to their individual dietary guidelines provided by the team's dietitian. Secondly, they rated their sleep quality and quantity to reflect how well refreshed athletes felt upon waking. Finally, a holistic "how you feel" score was recorded, indicative of an athletes' overall physical and psychological state, accounting for all aspects of perceptual wellbeing such as fatigue, mood, stress levels and soreness. In addition, six specific muscle sites were assessed for soreness (lower back, upper body, quadriceps, hamstrings, calves and groin), as they were identified by coaching staff to be of interest for their weekly player assessment.

INSERT FIGURE 1 NEAR HERE

Statistical Analyses

Descriptive statistics (mean ± SD) for weekly internal-TL and perceptual wellbeing ratings were calculated. Pearson's Chi-Square test was conducted to assess differences in the count of illness instances for each macrocycle. Prior to further analysis, internal-TL and perceptual wellness ratings were verified for the assumptions of homogeneity of variance. Comparisons of internal-TL and perceptual wellness ratings between each macrocycle were analysed using repeated measures analysis of variance (ANOVA), with a statistical significance level of $p < 0.05$. Following, a Bonferroni *post hoc* analysis was conducted. This analysis was performed using IBM Statistical Package for the Social Sciences (SPSS) (v22.0, IBM Corporation, Somers, New York, USA).

Beyond this, predictive models were developed to identify factors most likely to contribute to increased incidence of illness, including a decision tree, random forest and boosting models. These were developed using a range of variables from the dataset including training macrocycle, internal-TL values, wellbeing, muscle soreness ratings and the age of the athlete. In order to eliminate the variation of individual patterns in measures, raw data were converted to a Z-score. This represents the number of standard deviations the raw score is distributed from the mean of that particular athlete's previous data points. Additionally, to reflect a broader understanding of the overall muscle soreness experienced and wellbeing of the athlete, a mean value for each category was included. The models were further evaluated for their predictive abilities using an error matrix, which compares predictions with actual observations made. An error rate was also calculated as the proportion of incorrectly predicted outcomes by the total number of observations. Lastly, a receiver operating characteristic chart was included to provide information regarding the true positive rate against the false positive rate. The area under the curve for this chart was also examined, with a value of 1 indicating a 100% hit rate of predicting the target variable (illness). These analyses were completed using R (v R-3.1.3).[24] An additional graphical user interface package was used; Rattle (v3.4.2).[25] Both R and Rattle are open source software programs.

Results

Descriptive Statistics

Self-reported illness information, internal-TL, perceptual wellbeing and muscle soreness information for each macrocycle is reported in Table 2. Repeated measures ANOVA identified variation in internal-TL measures between macrocycles for weekly-TL ($F_{3,270} = 152.58$; $p < 0.001$; $\eta^2 = 0.63$), monotony ($F_{3,270} = 153.29$; $p < 0.001$; $\eta^2 = 0.63$) and strain ($F_{3,270} = 185.96$; $p < 0.001$; $\eta^2 = 0.67$). *Post hoc* analysis confirmed the largest difference in internal-TL measures were between specific preparation and competition phases for weekly-TL, monotony and strain, corresponding to a mean decrease of 2764 AU, 0.68 AU and 5191 AU, respectively ($p < 0.001$). Differences in weekly illness incidence between macrocycles were not significant, similarly observed for weekly wellbeing and muscle soreness ratings.

INSERT TABLE 2 NEAR HERE

Illness Prediction Models

Decision Tree Model

A decision tree model (Figure 2) was developed using a total of 556 observations from the dataset. The tree represents a series of decisions referred to as 'nodes' that are interpreted in a top-down manner. Each node is numbered for reference purposes (top left corner). The contributing variable is listed on the first line in each node, and the value of change associated with illness presented on the second line. The third line is the cross-validated error of the node, representing the associated change in the accuracy of the model as new levels are added to the tree. The model was appropriately pruned to reduce the complexity of the model and to improve the accuracy. To explain the interpretation of this model, root node 1 identified reductions below -2.05 SD of the 'how you feel' rating was identified as the first predictor contributing to the incidence of illness. If a 'yes' decision was made, the following node (node 2) recognized reductions in wellbeing average below 7.25 AU as a contributor. From here, the tree divided into two branches, whereby if a 'yes' decision was made, the following node (node 3) identified reductions in measures including food Z-score, how you feel Z-score and reductions in muscle soreness average, were all recognized as contributors. However, if a 'no' decision was made at node 2, weekly-TL in excess of 2765 AU was identified as a predictor of illness. Following this decision, node 6 showed increased strain values and monotony values to contribute to the incidence of illness. Overall, this model possessed an error of 5.9% determined by the error matrix, and using the ROC curve an area under the curve was calculated at 46%.

INSERT FIGURE 2 NEAR HERE

Random Forest Model

The random forest algorithm builds a single model based the combined information from a large number of decision trees. Predicting variables are ranked according to the number of times they are presented in the series of decision tree models developed. The random forest model is then further evaluated for its prediction accuracy (mean decrease accuracy) and the nodes impurity or splitting criterion (mean decrease Gini). This model (Figure 3) was developed using 556 observations, and included 500 trees with 3 variables tested at each split. This model identified internal-TL measures strain, weekly-TL and monotony as the three greatest contributors to the incidence of illness, followed by self-reported ratings for sleep and how you feel. This model possessed an overall error rate of 4.6% and an ROC of 74%.

INSERT FIGURE 3 NEAR HERE

Boosting Model

A boosting model was also developed (Figure 4), which provides information about contributing variables by taking into account the accuracy and inaccuracy of observations.[25] The boosting algorithm associates a weighting score with observations in the dataset, thus the final model is constructed from the series of models output of weighted scores of variables, ranked in order of importance. This model possessed an overall error of 5.2% and an ROC of 80%.

INSERT FIGURE 4 NEAR HERE

Discussion

The main findings of this investigation demonstrate that by analyzing self-report measures using statistical modelling techniques, it is possible to predict periods of increased illness risk for team-sport athletes. To our best knowledge, the present study is the first to utilize such predictive modelling techniques based on self-reported information in team-sport athletic cohorts. The results highlighted the influence of increasing training demands and fluctuations in athlete wellbeing between macrocycles as contributors to illness.

The training and match-play requirements for team-sports provide an environment that is likely to increase an athlete's exposure to pathogens at periods of heightened susceptibility to illness. It was not surprising to observe a total of 45 self-reported illness incidences during the data collection period, with the most commonly reported symptoms being runny nose, coughing and sore throat from the pre-determined symptoms. These symptoms may result in negative effects on performance and wellbeing as noted in previous research,[9] and may be indicative of excessive training stress. In the current study, ill athletes were often prevented from usual training and in the worst-case scenario prevented an affected athlete from match-play, which may potentially impact team performance.

Internal-TL values varied significantly according to macrocycle, reflecting the application of periodization principles. Mean monotony and strain values were greatest during the pre-season period particularly during specific preparation, and were comparable with values previously reported for professional rugby league athletes[1] and Australian football athletes.[7] The heightened TL during this period is necessary to maximise physical capacities required for competition. In contrast, during competition a significant reduction in TL was observed in order to optimize post-match recovery, where emphasis is on quality of training and peak match performance.[5]

The predictive models employed identified strain values >2282 AU as a significant contributor to illness. Previous research has recognized that 89% of illnesses could be explained by a preceding spike in strain of athletes from a wide range of competition levels.[22] More specifically, increased strain values has been identified to be associated with reduced salivary immunoglobulin A concentration and a "worse than normal" stress response on the Daily Analysis of Life Demands questionnaire.[26] The calculation of strain values for team-sport athletes is therefore useful as it reflects periods of intensified training with minimal recovery between sessions. Additionally, weekly-TL values >2786 AU were recognized as a contributing variable to self-reported illness incidence, ranking as the second greatest contributor in the random forest model. Monotony values >0.78 AU was noted in the decision tree model as a contributor to illness incidence, supporting previous research that identified a 10% increase in weekly-TL (to ~3,400 AU) and monotony (to ~1.19 AU) could explain a 42% and 33% increase in illness incidence, respectively.[7]

Collectively, the data presented in the current study provide evidence for the association between heightened internal-TL and the presence of self-reported illness. Whilst these data support previous research of the increased risk of illness associated with heightened training,[22,27,28] the results of the present study are specific to the athletic cohort recruited. Further, it must be noted that these findings are applicable to other team-sport athletes, however the training load values will vary depending on the typical training undergone by the athletes.

Wellbeing measures were recognized as predictors of illness in the series of predictive models produced, in particular reductions in athletes' average wellbeing and muscle soreness score. Variations in fatigue and psychological variables (e.g. mood and attitude) have been identified to be sensitive to training overload and cytokine markers, both known to have profound effects on immune functioning.[29] These data provide evidence for the use of subjective wellbeing questionnaires, given their ability to assess individual responses to exercise stress and recovery states,[2,18] thus, may be indicative of athletes' susceptibility to illness risk. Previous research has recognised that 80% of staff in high-performance sport have developed their own questionnaire, as they are seen as practical and effective for wellbeing assessment.[19]

While the current study provides important information regarding the use of self-report measures to predict illness, some limitations must be acknowledged. The present study did not assess pathological markers of illnesses, and solely relied on athletes' perceived symptoms. It is not often feasible to collect such markers of illness in the team-sport environment, given the logistical difficulties and associated costs, and therefore such analyses are not a common practice for similar athlete cohorts.[19] Further, it must be noted that an increase in support in the literature proposing that self-report measures may be more sensitive and reliable than traditional physiological, biochemical and performance measures is evident.[14,30,31]

Furthermore, the use self-report wellness information may be deemed a limitation as the reliability may be implicated if athletes report dishonest values. All athletes' were familiar with the questionnaire, as it had been used for numerous seasons. They were educated of the tasks and the importance of its use, therefore were encouraged to report honest values. Although athletes were required to report a perceived illness and the symptoms associated, no pathological infectious cause may be evident. Previous research has recognized athletes to be unable to distinguish between both infectious and non-infectious respiratory symptoms,[32] therefore for the present study this self-report method may be deemed as a limitation.

In conclusion, this study recognized the ability of self-report monitoring information to identify potentially contributing factors to the presence of self-reported illness for professional team-sport athletes. This was achieved by the development of a series of predictive models, using novel statistical modelling techniques. These findings emphasize the importance of multimodal athlete monitoring systems, incorporating the quantification of internal-TL and wellbeing responses given the predictive capacities these methods possess for the incidence of illness.

Practical Applications

Based on the current findings, it is evident that the internal-TL of team sport athletes should be monitored during all training sessions and matches, given that intensified internal-TL measures were found to predict self-reported illness. An important factor to consider is that monitoring markers of TL and athlete wellness may only be useful if they can be quickly analysed to highlight players who display early indications of an illness. If these tools are to be used successfully to identify an increased risk of developing an illness, they must be able to identify such players before they undergo additional exercise-related stress. A particular emphasis on this in the applied team-sport environment is prevalent, as informing best practice is key, possibly achieved using a predictive approach as opposed to retrospectively. Future research should also examine the integration of other monitoring methods (e.g. global positioning systems and accelerometer data) with self-report measures to highlight periods when athletes may be at increased risk of illness. This could provide a comprehensive monitoring strategy to identify players who may need to have training altered to limit the chance of developing an illness, or to promote recovery an already present illness.

Conclusion

The results of the present study provide important information regarding how contributing factors to illness can be monitored for team-sport athletes. More specifically, weekly-TL >2765 AU, monotony >0.78 AU and strain >2282 AU were strong predictors of the incidence of illnesses for the cohort of athletes. As such, coaching staff should aim to identify players who exhibit TL-related variables above these thresholds, as these individuals may be at increased risk of developing an illness. In addition, our data demonstrate that perceptual ratings of overall wellbeing and muscle soreness were also related to the incidence of self-reported illness. While wellbeing questionnaires are now commonplace in monitoring practices for high-level sport, the results of this investigation provide further evidence for the usefulness of these tools.

Acknowledgements

The authors would like to acknowledge the Newcastle Knights Rugby League Club for permitting this project to be conducted. Further, we wish to thank the athletes involved for their participation, and the coaching staff for their assistance. No financial assistance was provided for the current project. There were no conflicts of interest.

References

1. Coutts AJ, Sirotic AC, Catterick C, Knowles H. Monitoring training loads in professional rugby league. The Sixth World Congress on Science and Football; 2008; Antalya, Turkey.
2. Twist C, Highton J. Monitoring fatigue and recovery in rugby league players. *Int. J. Sports Physiol. Perform.* 2013;8(5):467-474.
3. Johnston RD, Gibson NV, Twist C, Gabbett TJ, MacNay SA, MacFarlane NG. Physiological responses to an intensified period of rugby league competition. *J Strength Cond Res.* 2013;27(3):643-654.
4. Sirotic AC, Coutts AJ, Knowles H, Catterick C. A comparison of match demands between elite and semi-elite rugby league competition. *J Sports Sci.* 2009;27(3):203-211.
5. Bompa T, Haff G. *Periodization; Theory and Methodology of Training.* Fifth Edition ed: Human Kinetics; 2009.
6. Gabbett T. The development and application of injury prediction model for noncontact, soft-tissue injuries in elite collision sport athletes. *J Strength Cond Res.* 2010;24(10):2593-2603.
7. Piggott B, Newton M, McGuigan M. The relationship between training load and incidence of injury and illness over a pre-season at an Australian football league club. *J Aust Strength Cond.* 2009;17(3):4-17.
8. Nieman DC. Exercise, upper respiratory tract infection, and the immune system. *Med Sci Sport Exercise.* 1994;26(2):128-139.
9. Nieman DC. Risk of upper respiratory tract infection in athletes: an epidemiologic and immunologic perspective. *J Athl Train.* 1997;32(4):344-349.
10. Spence L, Brown WJ, Pyne DB, et al. Incidence, etiology, and symptomatology of upper respiratory illness in elite athletes. *Med Sci Sport Exercise.* 2007;39(4):577-587.
11. Nieman DC, Bishop NC. Nutritional strategies to counter stress to the immune system in athletes, with special reference to football. *J Sports Sci.* 2006;24(07):763-772.
12. Weidner TG, Anderson BN, Kaminsky LA, Dick EC, Schurr T. Effect of a rhinovirus-caused upper respiratory illness on pulmonary function test and exercise responses. *Med Sci Sport Exercise.* 1997;29(5):604-609.
13. Coutts A, Reaburn P, Piva T, Rowsell G. Monitoring for overreaching in rugby league players. *Eur J Appl Physiol.* 2007;99:313-324.
14. Halson S. Monitoring training load to understand fatigue in athletes. *Sports Med.* 2014;44(2):139-147.
15. Stone MH, Stone M, Sands WA, Sands B. *Principles and practice of resistance training.* Human Kinetics; 2007.
16. McLellan CP, Lovell DI, Gass GC. Biochemical and endocine responses to impact and collision during elite rugby league match play. *J Strength Cond Res.* 2011;25(4):1030-1039.

17. Cormack SJ, Newton RU, McGuigan MR, Cormie P. Neuromuscular and endocrine responses of elite players during an Australian Rules Football Season. *Int. J. Sports Physiol. Perform.* 2008;3:439-453.
18. McLean BD, Coutts AJ, Kelly V, McGuigan MR, Cormack SJ. Neuromuscular, endocrine, and perceptual fatigue responses during different length between-match microcycles in professional rugby league players. *Int. J. Sports Physiol. Perform.* 2010;5(3):367-383.
19. Taylor K, Chapman D, Cronin J, Newton M, Gill N. Fatigue monitoring in high performance sport: A survey of current trends. *J Aust Strength Cond.* 2012;20(1):12-23.
20. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res.* 2001;15(1):109-115.
21. Rampinini E, Bishop D, Marcora S, M, Ferrari Bravo D, Sassi R, Impellizzeri F, M. Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *Int J Sports Med.* 2007;28(3):228-235.
22. Foster C. Monitoring training in athletes with reference to overtraining syndrome. *Med Sci Sport Exercise.* 1998;30(7):1164-1168.
23. Elloumi M, Makni E, Moalla W, et al. Monitoring training load and fatigue in rugby sevens players. *Asian J Sports Med.* 2012;3(3):175-184.
24. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria R Foundation for Statistical Computing; 2015.
25. Williams G. *Data mining with Rattle and R: The art of excavating data for knowledge discovery, use R!* : Springer; 2011.
26. Milanez VF, Ramos SP, Okuno NM, Boullosa DA, Nakamura FY. Evidence of a non-linear dose-response relationship between training load and stress markers in elite female futsal players. *J Spor Sci Med.* 2014;13:22-29.
27. Hausswirth C, Louis J, Aubry A, Bonnet G, Duffield R, Le Meur Y. Evidence of disturbed sleep and increased illness in overreached endurance athletes. *Med Sci Sport Exercise.* 2014;46(5):1036-1045.
28. Moreira A, de Moura NR, Coutts A, Costa EC, Kempton T, Aoki MS. Monitoring internal training load and mucosal immune responses in futsal athletes. *J Strength Cond Res.* 2013;27(5):1253-1259.
29. Main LC, Dawson B, Heel K, Grove RJ, Landers GJ, Goodman C. Relationship between inflammatory cytokines and self-report measures of training overload. *Res Sports Med.* 2010;18(2):127-139.
30. Buchheit M, Racinais S, Bilsborough J, et al. Monitoring fitness, fatigue and running performance during a pre-season

training camp in elite football players. *J Sci Med Sport.* 2013;16(6):550-555.

31. Coutts AJ, Slattery KM, Wallace LK. Practical tests for monitoring performance, fatigue and recovery in triathletes. *J Sci Med Sport.* 2007;10(6):372-381.

32. Cox AJ, Gleeson M, Pyne DB, Callister R, Hopkins WG, Fricker PA. Clinical and laboratory evaluation of upper respiratory symptoms in elite athletes. *Clin J Sports Med.* 2008;18(5):438-445.

Figure Captions

Figure 1. Subjective questionnaire used for the present study. Part A includes the self-reported incidence of illness and the symptoms experienced. Part B includes a muscle soreness and wellbeing component, scored using the 1-10 Likert rating scale as depicted.

Figure 2. Decision tree model representing a series of decisions referred to as 'nodes' that are numbered for reference purposes (top left of each node). The contributing variable is listed on the first line in each node, and the value of change associated with illness presented on the second line. The third line is the cross-validated error of the node, representing the associated change in the accuracy of the model as new levels are added to the tree.

Figure 3. Random forest model depicting predicting variables that are ranked according to the number of times it presents in the series of models developed. They are further evaluated for their prediction accuracy (mean decrease accuracy) and the nodes impurity or splitting criterion (mean decrease Gini).

Figure 4. Boosting model representing the variables of importance contributing to the incidence of illness. The boosting algorithm associates a weighting score with observations in the dataset, thus the final model is constructed from the series of models output of weighted scores of variables, ranked in order of importance.