Factors that impact self-reported wellness scores in elite Australian Footballers

Ruddy, Joshua D., Cormack, Stuart, Timmins, Ryan G., Sakadjian, Alex, Pietsch, Samuel, Carey, David L., Williams, Morgan D. and Opar, David A.

Title:

Factors that impact self-reported wellness scores in elite Australian footballers

Authors:

Joshua D. Ruddy[1], Stuart Cormack[1], Ryan G. Timmins[1], Alex Sakadjian[2], Samuel Pietsch[2], David L. Carey[3], Morgan D. Williams[4], David A. Opar[1]

[1]School of Behavioural and Health Sciences, Australian Catholic University, Melbourne, Australia

[2]Melbourne Football Club, Melbourne, Australia

[3]La Trobe Sport and Exercise Medicine Research Centre, College of Science, Health and Engineering, La Trobe University, Melbourne, Australia

[4]School of Health, Sport and Professional Practice, Faculty of Life Sciences and Education, University of South Wales, United Kingdom

Corresponding author:

Joshua D. Ruddy

+61 400 095 146

joshua.ruddy@myacu.edu.au

17 Young Street, Fitzroy, VIC, Australia (3065)

ABSTRACT

Introduction: This study aimed to 1) identify the impact of external load variables on changes in wellness and 2) identify the impact of age, training/playing history, strength levels and pre-season loads on changes in wellness in elite Australian footballers.

Methods: Data were collected from one team (45 athletes) during the 2017 season. Self-reported wellness was collected daily (4=best score possible, 28=worst score possible). External load/session availability variables were calculated using global positioning systems/session availability data from every training session and match. Additional variables included demographic data, pre-season external loads and strength/power measures. Linear mixed models were built and compared using root mean square error (RMSE) to determine the impact of variables on wellness.

Results: The external load variables explained wellness to a large degree (RMSE=1.55, 95% confidence intervals=1.52 to 1.57). Modelling athlete ID as a random effect appeared to have the largest impact on wellness, improving the RMSE by 1.06 points. Aside from athlete ID, the variable that had the largest (albeit negligible) impact on wellness was sprint distance covered across pre-season. Every additional 2.1 km covered across pre-season worsened athletes' in-season wellness scores by 1.2 points (95% confidence intervals=0.0 to 2.3).

Conclusion: The isolated impact of the individual variables on wellness was negligible. However, after accounting for the individual athlete variability, the external load variables examined collectively were were able to explain wellness to a large extent. These results validate the sensitivity of wellness to monitor individual athletes' responses to the external loads imposed on them.

Key words:

Australian football, athlete monitoring, wellness, training loads, mixed modelling

INTRODUCTION

Australian football is a team sport that requires a variety of skills, as well as large amounts of running, jumping and contact with opposition athletes (1). The external loads athletes are exposed to inevitably result in increased levels of fatigue following a match (2). High levels of fatigue are thought to increase the risk of subsequent injury via factors such as impaired neuromuscular control and tissue capacity (3). Additionally, greater levels of fatigue have been shown to negatively influence athletes' external load and performance during matches (2). As such, practitioners have a vested interest in monitoring athletes and their response to training and match demands.

Monitoring external training and match loads via global positioning systems (GPSs) is commonplace in team sports such as Australian football (4). The information provided to practitioners by GPS technology is often used to optimise training loads and ensure that athletes are ready to compete (4). However, it is also important to consider that external loads elicit different physiological and psychological responses (i.e. internal loads) in individual athletes (4). It is hypothesised that these individual responses are likely moderated by several other factors, such as age, playing/training history and fitness levels (3-5). Considering this, simply monitoring external loads may not inform practitioners as to how athletes are responding to training/match demands and their competition readiness. Accordingly, it is recommended that practitioners also implement methods to monitor athletes' internal responses.

Self-reported wellness questionnaires are a common method of monitoring athlete wellbeing, with a survey of practitioners working in high-performance sport reporting that 80% of responders implemented some form of customized questionnaire as part of their monitoring strategy (6). Additionally, subjective monitoring tools have been shown to respond to stress induced by training more consistently than objective measures (such as various hormonal/physiological markers) (7). Typically, self-reported wellness questionnaires focus on several different components (e.g. fatigue, sleep, soreness and stress), asking athletes to rate each component on a scale (7). It is commonplace to consolidate each component, or subscale, to indicate an athlete's overall wellness (8-10). Given the widespread application of self-reported wellness questionnaires, a number of studies have investigated the impact of training/competition on wellness and its various components. Studies in soccer, rugby and Australian football have reported declines in wellness in the days following matches (9, 11, 12). Additionally, self-reported wellness has been associated with subsequently modified external loads in elite soccer and Australian football athletes (8, 13) and is also suggested to influence the risk of future injury (7, 14).

Given the implications that wellness may have in regard to subsequent performance and injury risk, understanding the variables that impact wellness and the degree to which they do so may provide insights into the mechanisms responsible for changes in wellness. However, despite the prevalence of self-reported wellness questionnaires, limited research has investigated the impact of external loads on wellness in elite Australian footballers. Furthermore, despite previous research suggesting that responses to external loads may be moderated by individual athlete characteristics (3-5), no research has investigated the impact that such variables (beyond training/match loads) have on self-reported wellness. For example, playing experience and fitness/strength levels have been shown to moderate the impact of training/match loads on injury risk (15, 16). Whether the impact of training/match loads on wellness is moderated by individual athlete characteristics, however, is yet to be investigated.

Given the widespread application of wellness questionnaires in team sports (6), it is important to understand whether factors such as age, training/playing history and fitness/strength levels moderate the impact of training/match loads on self-reported wellness. This information may provide practitioners with a better understanding of the mechanisms responsible for observed differences in athletes' expected wellness scores versus their actual wellness scores. In turn, this may assist practitioners in making more meaningful inferences regarding athletes' responses to training/match demands and their competition readiness. Accordingly, the aims of the current study were to 1) identify the impact of external load variables on changes in self-reported wellness and 2) identify the impact of age, training/playing history, strength levels and pre-season loads on changes in self-reported wellness in elite Australian footballers.

METHODS

Study design

Data for this cohort study were collected during the 2017 Australian Football League (AFL) season (November 2016 to September 2017) and were obtained retrospectively by the research team. These data were collected from one team competing in the AFL. All athletes contracted to the team (n = 45) had their data included in this study (i.e. no athletes were excluded). This study was approved by the Australian Catholic University Human Research Ethics Committee (approval number: 2018-26WN).

Response variable

Throughout the in-season period (March 2017 to September 2017) athletes were instructed to complete a customised self-reported wellness questionnaire in private via an online system using their own device. Whilst the questionnaire used in the current study was customised, a similar questionnaire (with one additional component: mood) has been implemented in a number of prior studies (1, 8, 9, 12). The athletes completed the questionnaire on every morning they were at the football club prior to any training/activities and were not required to complete it on their days off or on match days. The questionnaire instructed athletes to rate their current level of fatigue, soreness, stress and sleep on a scale ranging from 1 (as good as possible) to 7 (as bad as possible). The sum of each subscale was then used to represent overall wellness, with a minimum score of 4 being the best possible and a maximum score of 28 being the worst possible. For every wellness measure, the number of days until the next AFL match was also determined, as this has previously been reported as the best predictor of wellness changes in elite Australian footballers (1).

Load variables

Athlete tracking data were collected for every field training session and match using valid (17) 10 Hz GPSs fitted into specially designed pockets on the back between the scapulae (OptimEye S5 GPS athlete monitoring systems, Catapult Sports, Melbourne, Australia). Using proprietary software (Openfield, Catapult Sports, Melbourne, Australia), the following data were extracted:

- Total distance – total distance (m) covered.
- High-speed running (HSR) distance – distance (m) covered above 16 km/h.
- Sprint distance – distance (m) covered above 26 km/h.

For each of these variables, exponentially weighted moving averages (EWMAs) were calculated for the day prior to each wellness measure being taken, using the following equation (18):

$$EWMA_{(current\ day)} = Load\ value_{(current\ day)} \times \lambda + \left((1-\lambda) \times EWMA_{(previous\ day)}\right)$$

$$\lambda = \frac{2}{N+1}$$

Where N is equal to the decay parameter. The decay parameter determines the weighting assigned to more recent and less recent observations, with a smaller decay parameter discounting less recent observations to a greater degree. EWMAs, as opposed to rolling averages, have been shown to provide a more sensitive marker of injury risk and are thought to better represent external loads (19). Several EWMAs were calculated using a 6-day and a 28-day decay parameter. Previous work has observed that a 6-day acute time window and a 28-day chronic time window best explained the risk of injury and it is suggested that these windows may be most appropriate for a typical microcycle in elite Australian football (20). Accordingly, the value of 6 was chosen to represent acute loads and the value of 28 was chosen to represent chronic loads. Using the EWMAs, a 6:28 day ratio was also calculated for each of the external load variables, where the chronic window was uncoupled from the acute window, as per the findings of previous work, which has shown that coupled acute and chronic windows can result in spurious correlations (21). These ratios were included in the analyses to determine whether they added any additional value beyond examining the acute and chronic loads as separate constructs.

In addition to the EWMAs and the ratios, each athlete's session availability was also determined using athlete participation data. Previous work has suggested that session availability may be a surrogate and potentially more accessible marker of load, compared to GPS/accelerometer derived variables (22). For every response measure, the number of training sessions and matches that were missed/modified (for any reason) in the prior 6, 28 and 84 days was determined for each athlete. The number of full training sessions and matches that each athlete could have conceivably completed was also determined for the same retrospective windows. Session availability (%) for each window was then determined as the number of training sessions and matches fully completed relative to the number of training sessions and matches possible for each athlete. The windows of 6 and 28 days were chosen as they correspond to the external load windows. However, the window of 84 days was chosen as previous work observed a significant interaction between injury risk and acute availability by availability in the prior 84 days (22). Session availability was examined to determine whether it could explain the response measures to the same degree as the external load EWMAs/ratios, as a potentially surrogate (and more easily accessible) marker of load.

Demographic, training/playing history and strength variables

Demographic data were collected at the beginning of the pre-season period (November 2016). These included date of birth (used to calculate age), stature (cm), mass (kg), years of AFL experience and the number of matches played in the prior season. A number of pre-season load variables were also determined, in addition to the aforementioned external load variables. Each athlete's total distance, HSR distance and sprint distance accumulated over pre-season were determined. Each athlete's session availability (%) over pre-season was also calculated. End of pre-season strength and power data were collected in February 2017. All strength and power data were collected using a 600 Hz force plate and analysed using proprietary software (Ballistic Measurement Software, Fitness Technologies, South Australia). Maximal isometric strength relative to mass (N/kg) was recorded using an isometric mid-thigh pull (IMTP) (23) and peak power relative to mass (W/kg) was recorded using a countermovement jump (CMJ) (24).

Due to a number of different reasons (such as an athlete being injured, ill or away at the time of testing) 14% of the strength and power measures were missing. One option to overcome the challenges of missing data is to exclude observations with missing data from the analyses. However, due to the limitations imposed on sports science/medicine researchers by small datasets (25), this option is undesirable. An alternative (and more pertinent) option is to replace the missing data via a process known as imputation (26). In the current study, multiple imputation by chained equations was implemented to replace the missing end of pre-season strength and power measures. Further details regarding the strength and power data collection methods and the imputation methods implemented in the current study can be found in Supplemental Digital Content 1.

STATISTICAL ANALYSIS

Prior to modelling the data, a correlation analysis was performed to identify redundant input variables. Reducing the number of input variables whilst retaining as much explanatory information as possible can improve the interpretation of a model and its coefficients (26). The correlation coefficient between each input variable was calculated. A Pearson's correlation coefficient threshold of > 0.80 was applied (26). If the pairwise correlation between two variables was > 0.80, the variable with the larger mean pairwise correlation (across all variables) was discarded, with the mean pairwise correlations being re-evaluated after the removal of every variable. A list of all the input variables prior to the correlation analysis can be found in Supplemental Digital Content 2.

Following the correlation analysis, models were constructed with the remaining input variables. These models have been detailed in Figure 1 and the R code corresponding to each of the models can be found in Supplemental Digital Content 3. The Akaike information criterion (AIC) of each model (constructed using the original data) was determined. AIC is a measure of the quality of a model that accounts for the trade-off between the complexity of the model (i.e. the number of input variables) and the fit of that model (i.e. how well the model can predict the response variable). A lower AIC indicates a better balance between model complexity and fit. Additionally, the average root mean square error (RSME) of each model was determined using 10-fold cross validation, repeated 10 times. The RMSE is equal to the standard deviation of the residuals (i.e. prediction errors) and is expressed on the same scale as the response variable. A lower RMSE indicates better predictive ability. The AICs were compared to determine which model provided the best balance between complexity and predictive ability, whilst the RMSEs were used to determine the absolute predictive ability of each model. The reader is directed to the following resource for further information regarding cross validation (25).

Following these comparisons, the coefficient and 95% confidence intervals (95% CIs) for each input variable were extracted from the model with the best predictive ability (i.e. lowest mean RMSE value) and interpreted. A coefficient was considered significant if the 95% CIs did not contain 0. All data/statistical analyses were performed using the R programming language (27) and the following packages: dplyr, caret, lme4 and ggplot2.

RESULTS

Cohort and descriptive details

Forty-five elite Australian footballers (age 24.0 ± 3.3 years, stature 188.0 ± 7.8 cm, mass 88.7 ± 7.9 kg and years of playing experience 4.3 ± 3.3) from one team competing in the AFL provided data for this study throughout the 2017 AFL season. Throughout the in-season period, a total of 3,267 wellness measures were collected, with each individual athlete providing, on average, 72.6 ± 15.7 measures. The mean self-reported wellness score was 15.7 points (Table 1). However, the within-individual mean self-reported wellness score ranged from 5.9 to 18.4 points, with the within-individual standard deviation ranging from 0.3 to 3.4 points. Descriptive statistics for each of the variables used to construct the models can be found in Table 1.

Correlation analysis

A full list of the input variables prior to the correlation analysis can be found in Supplemental Digital Content 2. A list of the remaining input variables following the correlation analysis can be found in Table 1. In total, 22 variables were reduced to 17 following the correlation analyses. The correlation matrix for all variables can be found in Supplemental Digital Content 4.

Model comparisons

The AIC and mean RMSE and 95% CIs of each model (calculated using repeated 10-fold cross validation) can be found in Table 2. Naïve Model 1 (NM1) had the best (i.e. lowest) AIC, suggesting that this model offered the best balance between complexity and predictive ability. However, the small AIC (relative to the other models) is likely due to NM1 only having one input variable (number of days until the next match). Whilst the introduction of athlete ID as random effect appeared to significantly increase the complexity of the remaining models (and subsequently the AIC), accounting for the individual athlete variability also reduced the RMSE and improved the predictive ability of the other models. The AICs for the remaining models were comparable.

In terms of absolute predictive ability, FM performed better than all other models, although the improvements in the RMSE were marginal, with FM improving on the performance of Load Model 3 (LM3) and Availability Model 3 (AM3) by 0.01 and 0.08 respectively. Load Model 2 (LM2) outperformed Availability Model 2 (AM2) by 0.07, suggesting that the session availability variables were not able to explain wellness to the same degree as the external load EWMAs/ratios. The inclusion of the demographic, training/playing history and strength variables did not improve the RMSE of LM3 and AM3 compared to LM2 and AM2 respectively. With the addition of athlete ID as a random effect, however, Naïve Model 2

(NM2) improved on the performance on Naïve Model 1 (NM1) by 1.06. Additionally, Load Model 1 (LM1) and Availability Model 1 (AM1), with the number of days until the next match excluded, were also compared to LM2 and AM2 respectively. LM2 (RMSE = 1.55) performed no better than LM1 (RMSE = 1.55). However, AM2 (RMSE = 1.62) outperformed AM1 by 0.02. Generally, the tight CIs suggest that the dataset is relatively homogenous, with variations in every iteration of the cross validated data having little effect on the RMSEs (Table 2).

Model coefficients

Athlete ID was modelled as a random effect and as such, its impact on wellness cannot be interpreted in the same manner as the fixed effects (i.e. other input variables). However, the largest improvement in RMSE was seen between NM1 and NM2 with the addition of athlete ID as a random effect. The conditional modes of the athletes are displayed in Figure 2. These conditional modes indicate the difference between the average (population-level) predicted wellness score and the predicted wellness score for the individual, for a given set of fixed input variables. Aside from athlete ID, the variable that had the largest individual impact on wellness was sprint distance covered across pre-season (km). An increase in sprint distance covered across pre-season equal to the interquartile range (2.1 km) increased (i.e. worsened) athletes' in-season wellness scores by 1.2 points (95% CIs = 0.0 to 2.3). The individual impact of all significant input variables on wellness has been illustrated in Figure 3. A full list of the coefficients can be found in Supplemental Digital Content 5.

DISCUSSION

Three of the major findings of the current study were 1) accounting for individual athlete variability had the largest impact on explaining changes in wellness, 2) the inclusion of the external load variables significantly improved the explanation of wellness and 3) the isolated impact of the individual external load variables on changes in wellness was negligible. It is important to consider that external loads will elicit different psychophysiological responses in individual athletes and that simply monitoring external loads may not inform practitioners as to how athletes are feeling and their competition readiness. Accordingly, self-reported wellness questionnaires are commonly implemented to monitor athlete wellbeing (6). However, despite the widespread application of self-reported wellness questionnaires, limited research has investigated the ability of wellness to capture the different psychophysiological responses elicited by external loads. In the current study, the within-individual mean wellness scores ranged from 5.9 to 18.4 points. Modelling athlete ID as a random effect, however, appeared to account for the different responses between athletes and had a large impact on the performance of the models.

The impact of modelling athlete ID as a random effect is further highlighted by the conditional modes illustrated in Figure 2. These conditional modes indicate the difference between the average (i.e. population-level) predicted wellness score and the predicted wellness score for the individual, for a given set of fixed input variables. For example, Athlete 45 can be expected to rate their wellness 8.9 points lower (i.e. better) than the average athlete, given the same fixed input variables, irrespective of their values. Athlete 1, however, can be expected to rate their wellness 3.2 points higher (i.e. worse) than the average athlete. The present findings highlight the importance of accounting for the variability amongst individual athletes when analysing and interpreting self-reported wellness data. A mixed modelling approach is an appropriate solution that accounts for individual variability and does not require the response variable (i.e. wellness) to be rescaled, which is important when determining the isolated impact of individual variables (as illustrated in Figure 3).

The RMSE of LM1 and LM2 was 1.55 (Table 1). Whilst there are no formal guidelines for interpreting RMSEs, given that the range of wellness in the current study is equal to 24, an RMSE of 1.55 suggests that the external load variables were able to explain wellness to a large extent. These findings are supported by the results of previous research, which attempted to predict the future wellness scores of elite soccer players using external load data and internal load (i.e. session ratings of perceived exertion [sRPE]) data (14). External load data improved the prediction of future fatigue, muscle soreness, mood and stress scores compared to a baseline model, which simply predicted each athlete's future score as their observed average (14). It was reported, however, that cumulative loads (i.e. loads in the prior 2, 3, 4 and 7 days), in addition to the previous day's training/match load, did not improve the prediction of future wellness scores (14). The authors of this study (14) suggest that external loads beyond those of the previous day are not meaningfully related to self-reported wellness. In contrast, the current study observed that the 28-day EWMA of HSR distance, in addition to the corresponding 6-day EWMA, significantly impacted wellness. It may be that the cumulative loads examined in previous work (14) (i.e. loads in the prior 2-7 days) were highly correlated with the previous day's load and that examining loads beyond the prior 7 days may have yielded different results.

Previous research has reported the number of days until the next match as the best predictor of self-reported wellness (1). In the current study, the number of days until the next match significantly impacted wellness, albeit to a small degree. Previous work has suggested that the association between the number of days until the next match and wellness may be the result of changes in external load due to the training/match schedule (9). Accordingly, LM1 was compared to LM2 to determine whether the inclusion of the number of days until the next match, beyond the external load variables, improved the explanation of wellness. Given that LM2 did not improve on the performance of LM1, it is possible that the number of days until the next match is simply a surrogate marker of external load. It is also important to consider that self-reported wellness is a perceptual tool and that athletes may actually 'feel' better in the lead up to the next match, without necessarily experiencing any improvements in a physiological sense (9). Alternatively, it may also be the case that athletes simply report better wellness in the lead up to the next match in order to improve their chances of selection, regardless of how they feel.

In addition to the number of days until the next match, HSR distances (defined as distances covered above 16 km/h) also significantly impacted wellness in the current study. Previous work in elite soccer has reported similar findings, with an association between distances covered above 14 km/h and self-reported fatigue being observed (28, 29). Presently, an increase in athletes' 6-day EWMA of HSR distance resulted in a higher (i.e. worse) wellness score (Figure 3C). However, it should be noted that a higher 28-day EWMA of HSR distance actually improved athletes' wellness scores (Figure 3D). Previous work suggests that accumulating higher chronic loads results in increased fitness levels and an increased tolerance of higher acute loads, which may explain the present findings (3). Despite these suggestions, however, the isolated impact of the individual external load variables on wellness was negligible. Decreasing an athlete's 6-day EWMA of HSR distance by 388.9 m (equal to the interquartile range) only improved their wellness score by 1.1 points (95% CIs = 1.0 to 1.2) (Figure 3C). During the in-season period, the mean number of training sessions/matches per week was 3 ± 1 and the mean weekly HSR distance covered by each individual athlete was 3957 ± 2197 m. Considering this, for an athlete that covers 3957 m of HSR across 3 training sessions/matches per week, reducing their 6-day EWMA of HSR distance by 388.9 m would require a decrease in HSR distance of approximately 878 m per session. Put simply, improving an athlete's wellness score by as little as 1 point would require large decreases in HSR distances. As per previous research (14), the current results suggest that changes in wellness are likely a function of complex, non-linear interactions amongst multiple variables and that targeting and modifying one specific variable (e.g. HSR distance) is unlikely to have any substantial impact.

There are a number of potential limitations in the current study. Firstly, internal load (i.e. sRPE) data were not available. However, the evidence regarding the impact of internal load on self-reported wellness is conflicting. One study, conducted in elite Australian football, observed an association between changes in sRPE-derived training load and changes in self-reported wellness (30). In contrast, other research has reported that self-reported wellness was not sensitive to changes in sRPE-derived training load (9, 14). Given the conflicting evidence, it is difficult to determine whether the inclusion of internal load data in the current analyses would have impacted the results. Secondly, the self-reported wellness questionnaire implemented in the current study was customised and has not been previously investigated. However, a similar questionnaire (with one additional component: mood) has been implemented in a number of prior studies (1, 8, 9, 12). Lastly, the sum of each subscale was used to represent overall wellness in the current study. Whilst consolidating the subscales is a common approach in practice and has been investigated in several prior studies (8-10), a systematic review has suggested that the consolidation of subscales into an overall score reduces the sensitivity of the measures (7). Despite this, all four of the subscales utilised in the current study have been previously investigated and reported as responsive to training/match demands (7). Additionally, subscales are typically measure on a Likert scale and should be treated as an ordinal response when examined individually. Previous research, however, has treated these ordinal responses as continuous data, which can have methodical implications (1, 14). Given this, future research should carefully consider the methods implemented to analyse and interpret self-reported wellness data.

Despite the aforementioned limitations, external loads appear to have a large impact on self-reported wellness. Given this, the results of the current study validate the sensitivity of wellness (as determined in this study) to monitor individual athletes' responses to the external loads imposed on them. However, the current results also highlight the importance of accounting for the variability amongst individual athletes when analysing and interpreting wellness data. The present findings also support existing evidence that suggests HSR distances significantly impact wellness (and its components) (28, 29). Despite this, the isolated, individual impact of the external load variables on wellness was negligible and modifying one specific variable (e.g. HSR distance) is unlikely to have any substantial effect. However, it should be noted that there may exist other individual variables, beyond those examined in the current study, that provide a more global indication of external load and impact wellness to a larger degree. Nonetheless, as per previous work (14), changes in wellness appear to be a function of complex, non-linear interactions amongst multiple variables and the current results support the need for an individualised, multifaceted approach to athlete monitoring. Previous research has suggested that implementing such an approach to monitor and compare athletes' expected wellness scores versus their actual wellness scores may assist practitioners in their load management strategies (14). Further research is needed, however, to determine the impact that this information may have on subsequent performance and injury risk.

In conclusion, the current study observed that accounting for individual athlete variability had the largest impact on self-reported wellness. Additionally, despite the negligible impact of the individual variables, the external load variables examined collectively were able to explain wellness to a large extent. The present findings validate the sensitivity of wellness to monitor individual athletes' responses to the external loads they are exposed to. Implementing such an approach may provide further insights into the mechanisms responsible for changes in wellness and may assist practitioners in using wellness data to make meaningful inferences regarding athletes' responses to training/match demands and their competition readiness.

Figure captions:

Figure 1. The variables used to construct each of the mixed models. The Akaike information criterion and the average root mean square error of each model was determined. Following this, a number of comparisons between the models were made to determine whether the inclusion/exclusion of different input variables explained wellness to differing degrees. EWMA, exponentially weighted moving average.

Figure 2. The conditional modes of the athletes, as a result of modelling athlete ID as a random effect. The conditional modes indicate the difference between the average (i.e. population-level) predicted wellness score, represented by the vertical dashed line, and the predicted wellness score for the individual, for a given set of fixed input variables. For example, Athlete 45 can be expected to rate their wellness 8.9 points lower (i.e. better) than the average athlete, given the same fixed input variables, irrespective of their values. Athlete 1, however, can be expected to rate their wellness 3.2 points higher (i.e. worse) than the average athlete.

Figure 3. The individual impact of the significant input variables on wellness (whilst all other input variables are held at their observed means). The y-axis indicates the change in wellness (after accounting for the conditional modes illustrated in Figure 2). The horizontal dashed line represents no change in wellness. The 95% confidence intervals are indicated by the grey shaded areas. The reader should note that a negative change indicates a better wellness score, whereas a positive change indicates a worse wellness score. EWMA, exponentially weighted moving average.

Declaration:

The authors declare that the results of the study are presented clearly, honestly, and without fabrication, falsification, or inappropriate data manipulation. The results of this study do not constitute endorsement by ACSM.

Competing interests:

The authors declare no competing interests.

Contributorship:

JR, AS and SP contributed to the design of the study. JR, AS, SP and SC contributed to the collection of the data. JR performed the data analysis. JR, DC and MW performed the statistical analysis. JR and DO drafted the manuscript. AS, SP, DC, MW, SC and RT contributed to the manuscript.

Funding:

No funding was received.

References:

1. 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.

2. Mooney MG, Cormack S, O'Brien BJ, Morgan WM, McGuigan M. Impact of neuromuscular fatigue on match exercise intensity and performance in elite australian football. J Strength Cond Res. 2013;27(1):166-173.
3. Windt J, Gabbett TJ. How do training and competition workloads relate to injury? The workload-injury aetiology model. Br J Sports Med. 2017;51(5):428-435.
4. Bourdon PC, Cardinale M, Murray A et al. Monitoring athlete training loads: consensus statement. Int J Sports Physiol Perform. 2017;12:2161-2170.
5. Nedelec M, McCall A, Carling C, Legall F, Berthoin S, Dupont G. The influence of soccer playing actions on the recovery kinetics after a soccer match. J Strength Cond Res. 2014;28(6):1517-1523.
6. Taylor K, Chapman D, Cronin J, Newton MJ, Gill N. Fatigue monitoring in high performance sport: a survey of current trends. J Aust Strength Cond. 2012;20(1):12-23.
7. Saw AE, Main LC, Gastin PB. Monitoring the athlete training response: Subjective self-reported measures trump commonly used objective measures: a systematic review. Br J Sports Med. 2016;50(5):281-291.
8. Gallo TF, Cormack SJ, Gabbett TJ, Lorenzen CH. Pre-training perceived wellness impacts training output in australian football players. J Sports Sci. 2016;34(15):1445-1451.
9. Gallo TF, Cormack SJ, Gabbett TJ, Lorenzen CH. Self-reported wellness profiles of professional australian football players during the competition phase of the season. J Strength Cond Res. 2017;31(2):495-502.
10. Sampson JA, Murray A, Williams S, Sullivan A, Fullagar HHK. Subjective wellness, acute:chronic workloads, and injury risk in college football. J Strength Cond Res. 2019; doi: 10.1519/JSC.0000000000003000.

11. Thorpe RT, Strudwick AJ, Buchheit M, Atkinson G, Drust B, Gregson W. Tracking morning fatigue status across in-season training weeks in elite soccer players. Int J Sports Physiol Perform. 2016;11(7):947-952.

12. 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.

13. Malone S, Owen A, Newton M et al. Wellbeing perception and the impact on external training output among elite soccer players. J Sci Med Sport. 2018;21(1):29-34.

14. Jaspers A, De Beeck TO, Brink MS et al. Predicting future perceived wellness in professional soccer: the role of preceding load and wellness. Int J Sports Physiol Perform. 2019;14(8):1074-1080.

15. Malone S, Hughes B, Doran DA, Collins K, Gabbett TJ. Can the workload–injury relationship be moderated by improved strength, speed and repeated-sprint qualities? J Sci Med Sport. 2018;22(1):29-34.

16. Malone S, Roe M, Doran DA, Gabbett TJ, Collins KD. Aerobic fitness and playing experience protect against spikes in workload: the role of the acute:chronic workload ratio on injury risk in elite gaelic football. Int J Sports Physiol Perform. 2017;12(3):393-401.

17. Johnston RJ, Watsford ML, Kelly SJ, Pine MJ, Spurrs RW. Validity and interunit reliability of 10 Hz and 15 Hz GPS units for assessing athlete movement demands. J Strength Cond Res. 2014;28(6):1649-1655.

18. Williams S, West S, Cross MJ, Stokes KA. Better way to determine the acute:chronic workload ratio? Br J Sports Med. 2017;51(3):209-210.

19. Murray NB, Gabbett TJ, Townshend AD, Blanch P. Calculating acute:chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. Br J Sports Med. 2017;51(9):749-754.

20. Carey DL, Blanch P, Ong KL, Crossley KM, Crow J, Morris ME. Training loads and injury risk in australian football-differing acute:chronic workload ratios influence match injury risk. Br J Sports Med. 2017;51(16):1251-1220.

21. Lolli L, Batterham AM, Hawkins R et al. Mathematical coupling causes spurious correlation within the conventional acute-to-chronic workload ratio calculations. 2019;53(15):921-922.

22. Ruddy JD, Pietsch S, Maniar N et al. Session availability as a result of prior injury impacts the risk of subsequent non-contact lower limb injury in elite Australian footballers. Front Physiol. 2019; https://doi.org/10.3389/fphys.2019.00737.

23. West DJ, Owen NJ, Jones MR et al. Relationships between force-time characteristics of the isometric midthigh pull and dynamic performance in professional rugby league players. J Strength Cond Res. 2011;25(11):3070-3075.

24. Cormack SJ, Newton RU, McGuigan MR, Doyle TL. Reliability of measures obtained during single and repeated countermovement jumps. Int J Sports Physiol Perform. 2008;3(2):131-144.

25. Ruddy JD, Cormack SJ, Whiteley R, Williams MD, Timmins RG, Opar DA. Modeling the risk of team sport injuries: a narrative review of different statistical approaches. Front Physiol. 2019; https://doi.org/10.3389/fphys.2019.00829.

26. Harrell Jr FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. 2nd ed. New York (NY): Springer; 2015.

27. R Core Team. R: A language and envrionement for statistical computing. 2013; Vienna: R Foundation for Statistical Computing.

28. Thorpe RT, Strudwick AJ, Buchheit M, Atkinson G, Drust B, Gregson W. Monitoring fatigue during the in-season competitive phase in elite soccer players. Int J Sports Physiol Perform. 2015;10(8):958-964.

29. Thorpe RT, Strudwick AJ, Buchheit M, Atkinson G, Drust B, Gregson W. The influence of changes in acute training load on daily sensitivity of morning-measured fatigue variables in elite soccer players. Int J Sports Physiol Perform. 2017;12:2107-2113.

30. Buchheit M, Racinais S, Bilsborough JC 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.

Supplemental digital content:

Supplemental Digital Content 1. docx

Extended methodology for the strength and power data collection and imputation.

Supplemental Digital Content 2. tif

A list of the input variables prior to the correlation analysis and the input variables following the correlation analysis. The remaining input variables were used to construct a series of models, as outlined in the methods section of the paper.

Supplemental Digital Content 3. docx

The R code used to construct each of the models.

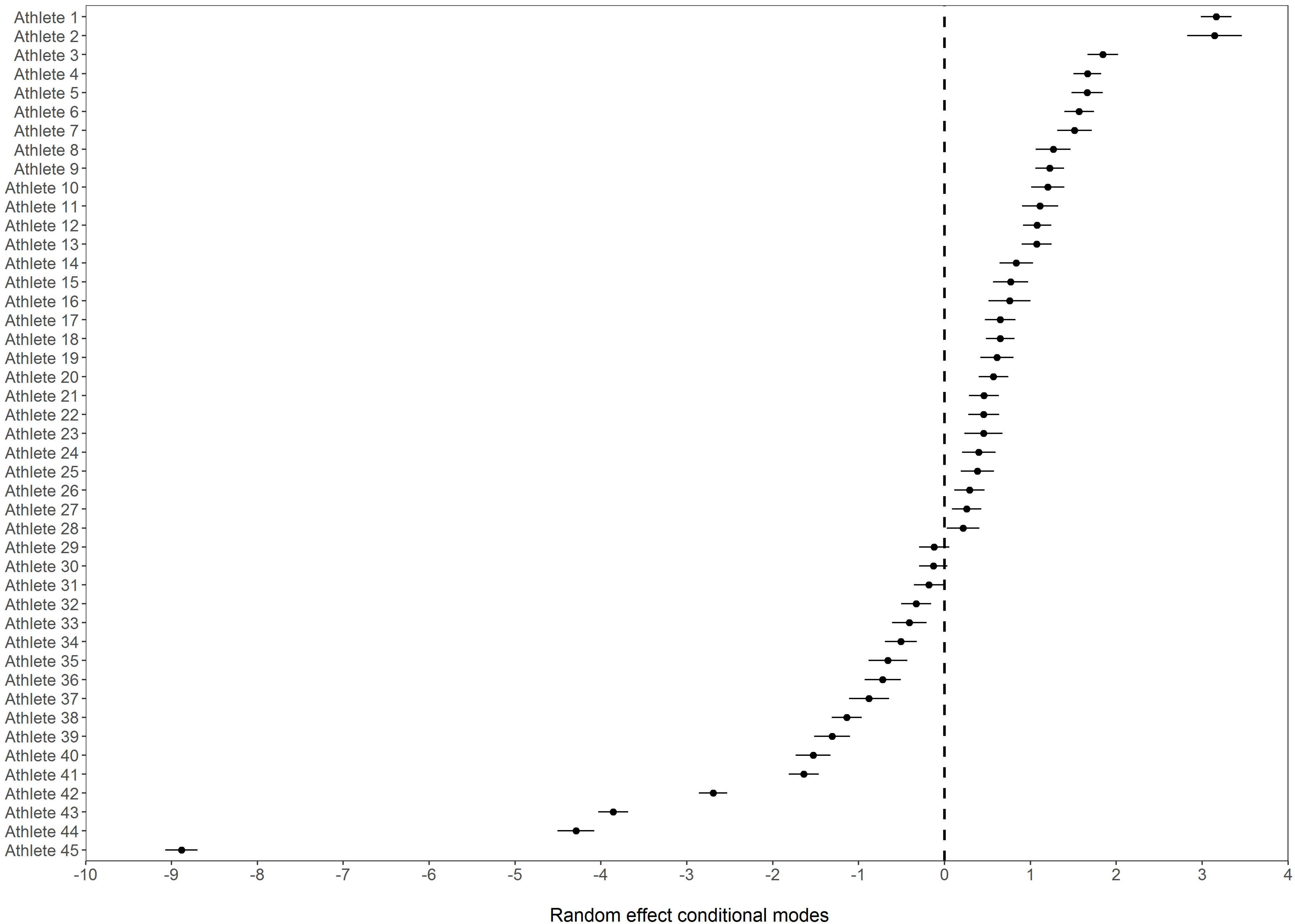
Supplemental Digital Content 4. docx

The pairwise correlation coefficient between all input variables.

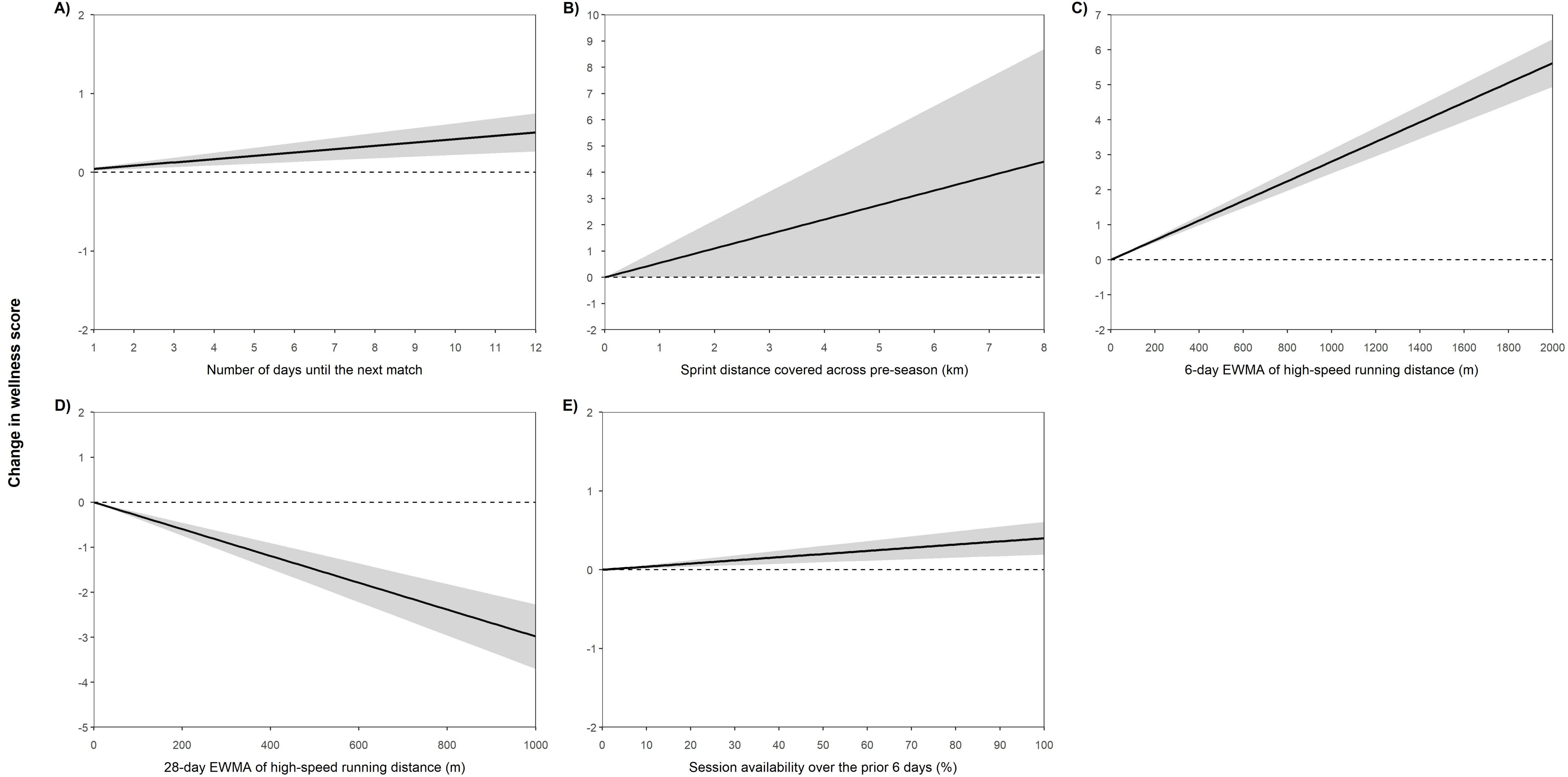
Supplemental Digital Content 5. docx

The coefficient and 95% confidence intervals (95% CIs) for all input variables, extracted from Full Model.

Variable	Naïve Model 1 (NM1)	Naïve Model 2 (NM2)	Load Model 1 (LM1)	Load Model 2 (LM2)	Load Model 3 (LM3)	Availability Model 1 (AM1)	Availability Model 2 (AM2)	Availability Model 3 (AM3)	Full Model (FM)
Wellness score (response variable)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of days until the next match	✓	✓		✓	✓		✓	✓	✓
Athlete ID (random effect)		✓	✓	✓	✓	✓	✓	✓	✓
Age					✓			✓	✓
Number of matches played in the prior season					✓			✓	✓
Isometric mid-thigh pull peak force					✓			✓	✓
Countermovement jump peak power					✓			✓	✓
Total distance across pre-season					✓			✓	✓
Sprint distance across pre-season					✓			✓	✓
Session availability across pre-season					✓			✓	✓
6-day EWMA of high-speed running distance			✓	✓	✓				✓
6-day EWMA of sprint distance			✓	✓	✓				✓
28-day EWMA of high-speed running distance			✓	✓	✓				✓
6:28 day ratio of total distance			✓	✓	✓				✓
6:28 day ratio of high-speed running distance			✓	✓	✓				✓
6:28 day ratio of sprint distance			✓	✓	✓				✓
Session availability in the prior 6 days						✓	✓	✓	✓
Session availability in the prior 84 days						✓	✓	✓	✓



Athlete 1
Athlete 2
Athlete 3
Athlete 4
Athlete 5
Athlete 6
Athlete 7
Athlete 8
Athlete 9
Athlete 10
Athlete 11
Athlete 12
Athlete 13
Athlete 14
Athlete 15
Athlete 16
Athlete 17
Athlete 18
Athlete 19
Athlete 20
Athlete 21
Athlete 22
Athlete 23
Athlete 24
Athlete 25
Athlete 26
Athlete 27
Athlete 28
Athlete 29
Athlete 30
Athlete 31
Athlete 32
Athlete 33
Athlete 34
Athlete 35
Athlete 36
Athlete 37
Athlete 38
Athlete 39
Athlete 40
Athlete 41
Athlete 42
Athlete 43
Athlete 44
Athlete 45
-10
-9
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
Random effect conditional modes



A)
B)
C)
D)
E)
Change in wellness score
Number of days until the next match
Sprint distance covered across pre-season (km)
6-day EWMA of high-speed running distance (m)
28-day EWMA of high-speed running distance (m)
Session availability over the prior 6 days (%)

Table 1. Descriptive statistics for the variables remaining following the correlation analysis.

Variable	Mean	Median	Standard deviation	25th percentile	75th percentile	Interquartile range
Wellness score	15.7	16	2.7	15	17	2
Number of days until the next match	4.0	4.0	2.8	2.0	5.0	3.0
Age (years)	24.0	23.2	3.3	21.4	26.0	4.6
Number of matches played in the prior season	15.8	19.0	7.5	14.0	21.0	7.0
Isometric mid-thigh pull peak force (N/kg)	39.8	38.8	5.4	37.6	42.3	4.7
Countermovement jump peak power (W/kg)	53.2	52.5	5.9	49.8	55.9	6.1
Total distance across pre-season (km)	350.0	351.3	41.5	324.8	378.6	53.8
Sprint distance across pre-season (km)	4.0	3.7	1.5	3.0	5.1	2.1
Session availability across pre-season (%)	76.1	80.4	17.7	61.7	89.1	27.4
6-day EWMA of high-speed running distance (m)	597.0	566.0	319.7	396.2	785.1	388.9
6-day EWMA of sprint distance (m)	24.3	19.1	22.3	8.2	35.0	26.8
28-day EWMA of high-speed running distance (m)	607.7	635.2	195.7	529.0	729.8	200.8
6:28 day ratio of total distance	1.1	0.9	3.3	0.7	1.2	0.5
6:28 day ratio of high-speed running distance	1.1	0.9	2.7	0.7	1.3	0.6
6:28 day ratio of sprint distance	2.8	0.7	42.2	0.4	1.3	0.9
Session availability in the prior 6 days (%)	72.6	100.0	38.5	50.0	100.0	50.0
Session availability in the prior 84 days (%)	72.3	82.4	25.4	60.0	91.4	31.4

EWMA, exponentially weighted moving average

Table 2. The Akaike information criterion (AIC) for each model and the average root mean square error (RMSE) and 95% confidence intervals (95% CIs) for each model. The RMSE and the 95% CIs were calculated using repeated 10-fold cross validation. The input variables used to construct each of the models are illustrated in Figure 1.

		RMSE	
Model	AIC	Mean	95% CIs
Naïve Model 1 (NM1)	5873	2.70	2.67 to 2.74
Naïve Model 2 (NM2)	11443	1.64	1.62 to 1.65
Load Model 1 (LM1)	11102	1.55	1.53 to 1.57
Load Model 2 (LM2)	11088	1.55	1.52 to 1.57
Load Model 3 (LM3)	11094	1.55	1.52 to 1.57
Availability Model 1 (AM1)	11457	1.64	1.62 to 1.66
Availability Model 2 (AM2)	11397	1.62	1.60 to 1.64
Availability Model 3 (AM3)	11403	1.62	1.60 to 1.64
Full Model (FM)	11074	1.54	1.52 to 1.56

Supplemental Digital Content 1. Extended methodology for the strength and power data collection and imputation.

Start and end of pre-season strength and power data were collected in November 2016 and February 2017 respectively. All strength and power data were collected using a 600 Hz force plate and analysed using proprietary software (Ballistic Measurement Software, Fitness Technologies, South Australia).

Maximal isometric strength relative to mass (N/kg) was recorded using an isometric mid-thigh pull (IMTP) [1]. Athletes stood on the force plate and held an immovable bar, using wrist straps to assist their grip [1]. The bar was fixed at an individualised height for each athlete that allowed for a hip angle of approximately 155-165 degrees and a knee angle of approximately 125-135 degrees [1]. Athletes were instructed to pull up as hard and as fast as possible for approximately five seconds [1]. Following a warm-up (self-perceived 75% of each athlete's maximum), only one maximum IMTP trial was performed.

Peak power relative to mass (W/kg) was recorded using a countermovement jump (CMJ) [2]. Athletes stood on the force plate and were instructed to maintain their hands on their hips throughout the jump and jump as high as possible [2]. Following warm-up (self-perceived 75% of each athlete's maximum), only one maximum CMJ trial was performed.

Due to a number of different reasons (such as an athlete being injured, ill or away at the time of testing) 14% of the end of pre-season strength and power measures were missing. One option to overcome the challenges of missing data is to exclude observations with missing data from the analyses. However, due to the limitations imposed on sports science/medicine researchers by small datasets [3], this option is undesirable. An alternative (and more pertinent) option is to replace the missing data via a process known as imputation [4]. In the current study, multiple imputation by chained equations was implemented.

Prior to imputing the missing data, a stepwise approach was implemented to determine which variables were best suited to impute the missing data. Out of observations with known strength and power values, 15% were withheld as a testing set. The withheld strength and power values of the testing set were imputed using the remaining 85% of observations. The imputed (i.e. predicted) strength and power values were then compared to the actual strength and power values and the root mean square error (RMSE) was calculated. This process was repeated, with a different variable being removed each iteration until no further variables could be removed without an increase in the RMSE. The variables that best predicted the strength and power values of the withheld testing set were mass (kg), all remaining IMTP peak force (N/kg) values and all remaining CMJ peak power (W/kg) values. One thousand iterations of this process were performed, resulting in a mean RMSE of 4.5 (95% confidence intervals = 4.4 to 4.6).

Following this process, the original missing end of pre-season strength and power measures were imputed using mass (kg) and all remaining start and end of pre-season strength and power measures. Fifteen imputations were performed over 50 iterations. For each missing data point, the mean of all its imputed value was used as the final prediction. The final predicted value was then used for the analyses outline in the methods section of the paper.

References (for Supplementary Material 1)

1. West DJ, Owen NJ, Jones MR, Bracken RM, Cook CJ, Cunningham DJ, Shearer DA, Finn CV, Newton RU, Crewther BT, Kilduff LP. Relationships between force-time characteristics of the isometric midthigh pull and dynamic performance in professional rugby league players. J Strength Cond Res. 2011;25(11):3070-3075.

2. Cormack SJ, Newton RU, McGuigan MR, Doyle TLJIjosp, performance. Reliability of measures obtained during single and repeated countermovement jumps. Int J Sports Physiol Perform. 2008;3(2):131-144.

3. Ruddy JD, Cormack SJ, Whiteley R, Williams MD, Timmins RG, Opar DA. Modeling the risk of team sport injuries: a narrative review of different statistical approaches. Front Physiol. 2019; https://doi.org/10.3389/fphys.2019.00829.

4. Harrell Jr FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. 2015; New York: Springer.

1. West, D.J., et al., *Relationships between force-time characteristics of the isometric midthigh pull and dynamic performance in professional rugby league players.* J Strength Cond Res, 2011. **25**(11): p. 3070-5.

2. Cormack, S.J., et al., *Reliability of Measures Obtained During Single and Repeated Countermovement Jumps.* International Journal of Sports Physiology and Performance, 2008. **3**(2): p. 131-144.

3. Ruddy, J.D., et al., *Modeling the Risk of Team Sport Injuries: A Narrative Review of Different Statistical Approaches.* Front Physiol, 2019. **10**: p. 829.

4. Harrell Jr, F.E., *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. 2015: Springer.

Variable	Remaining after the correlation analysis
Athlete ID	✓
Number of days until the next match	✓
Age (years)	✓
Number of matches played in the prior season	✓
Isometric mid-thigh pull peak force (N/kg)	✓
Countermovement jump peak power (W/kg)	✓
Total distance across pre-season (km)	✓
High-speed running distance across pre-season (km)	
Sprint distance across pre-season (km)	✓
Session availability across pre-season (%)	✓
6-day EWMA of total distance (m)	
6-day EWMA of high-speed running distance (m)	✓
6-day EWMA of sprint distance (m)	✓
28-day EWMA of total distance (m)	
28-day EWMA of high-speed running distance (m)	✓
28-day EWMA of sprint distance (m)	
6:28 day ratio of total distance	✓
6:28 day ratio of high-speed running distance	✓
6:28 day ratio of sprint distance	✓
Session availability in the prior 6 days (%)	✓
Session availability in the prior 28 days (%)	
Session availability in the prior 84 days (%)	✓

Supplemental Digital Content 2. A list of all input variables prior to the correlation analysis and the input variables remaining following the correlation analysis. The remaining input variables were used to construct a series of models, as outlined in the methods section of the paper.

Supplemental Digital Content 3. The R code used to construct each of the models.

```
library(lme4)

nm1 <- lm(wellness ~ days_until_next_match,
          data = train_data)

nm2 <- lmer(wellness ~ days_until_next_match + (1|id),
            data = train_data, REML = F)

lm1 <- lmer(wellness ~ (1|id) +
              HSR_distance_6_day_EWMA +
              HSR_distance_28_day_EWMA +
              sprint_distance_6_day_EWMA +
              total_distance_6to28_ratio +
              HSR_distance_6to28_ratio +
              sprint_distance_6to28_ratio,
            data = train_data, REML = F)

lm2 <- lmer(wellness ~ days_until_next_match + (1|id) +
              HSR_distance_6_day_EWMA +
              HSR_distance_28_day_EWMA +
              sprint_distance_6_day_EWMA +
              total_distance_6to28_ratio +
              HSR_distance_6to28_ratio +
              sprint_distance_6to28_ratio,
            data = train_data, REML = F)
```

```
lm3 <- lmer(wellness ~ days_until_next_match + (1|id) +
              HSR_distance_6_day_EWMA +
              HSR_distance_28_day_EWMA +
              sprint_distance_6_day_EWMA +
              total_distance_6to28_ratio +
              HSR_distance_6to28_ratio +
              sprint_distance_6to28_ratio +
              age +
              matches_played_prior_season +
              preseason_availability +
              preseason_total_distance +
              preseason_sprint_distance +
              imtp +
              cmj,
            data = train_data, REML = F)

am1 <- lmer(wellness ~ (1|id) +
              availability_6_day +
              availability_84_day,
            data = train_data, REML = F)

am2 <- lmer(wellness ~ days_until_next_match + (1|id) +
              availability_6_day +
              availability_84_day,
            data = train_data, REML = F)
```

```
am3 <- lmer(wellness ~ days_until_next_match + (1|id) +
              availability_6_day +
              availability_84_day +
              age +
              matches_played_prior_season +
              preseason_availability +
              preseason_total_distance +
              preseason_sprint_distance +
              imtp +
              cmj,
            data = train_data, REML = F)

fm <- lmer(wellness ~ days_until_next_match + (1|id) +
             HSR_distance_6_day_EWMA +
             HSR_distance_28_day_EWMA +
             sprint_distance_6_day_EWMA +
             total_distance_6to28_ratio +
             HSR_distance_6to28_ratio +
             sprint_distance_6to28_ratio +
             age +
             matches_played_prior_season +
             preseason_availability +
             preseason_total_distance +
             preseason_sprint_distance +
             imtp +
             cmj +
             availability_6_day +
             availability_84_day,
           data = train_data, REML = F)
```

Supplemental Digital Content 4. The pairwise correlation coefficient between all input variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	-	0.09	0.12	0.00	0.03	0.04	0.06	0.28	0.20	0.12	0.18	0.05	0.19	0.08	0.24	0.25	0.04	0.06	0.02	0.02	0.01	0.01
2	0.09	-	0.02	0.03	0.00	0.01	0.01	0.01	0.02	0.00	0.12	0.06	0.18	0.09	0.05	0.04	0.00	0.00	0.01	0.02	0.01	0.01
3	0.12	0.02	-	0.21	0.01	0.15	0.24	0.16	0.00	0.15	0.01	0.00	0.05	0.05	0.01	0.03	0.04	0.04	0.03	0.02	0.03	0.05
4	0.00	0.03	0.21	-	0.31	0.37	0.31	0.26	0.09	0.21	0.07	0.10	0.03	0.04	0.01	0.02	0.02	0.02	0.02	0.08	0.10	0.13
5	0.03	0.00	0.01	0.31	-	0.48	0.35	0.11	0.01	0.04	0.03	0.05	0.02	0.03	0.07	0.09	0.00	0.01	0.03	0.03	0.00	0.12
6	0.04	0.01	0.15	0.37	0.48	-	0.86	0.43	0.00	0.27	0.14	0.24	0.12	0.19	0.03	0.06	0.01	0.01	0.01	0.11	0.15	0.21
7	0.06	0.01	0.24	0.31	0.35	0.86	-	0.66	0.20	0.14	0.10	0.18	0.14	0.22	0.12	0.20	0.00	0.00	0.00	0.08	0.10	0.13
8	0.28	0.01	0.16	0.26	0.11	0.43	0.66	-	0.42	0.21	0.04	0.06	0.12	0.19	0.35	0.51	0.02	0.02	0.01	0.05	0.06	0.08
9	0.20	0.02	0.00	0.09	0.01	0.00	0.20	0.42	-	0.27	0.07	0.11	0.04	0.06	0.04	0.08	0.02	0.03	0.04	0.09	0.10	0.08
10	0.12	0.00	0.15	0.21	0.04	0.27	0.14	0.21	0.27	-	0.17	0.28	0.15	0.23	0.09	0.14	0.01	0.02	0.00	0.20	0.23	0.23
11	0.18	0.12	0.01	0.07	0.03	0.14	0.10	0.04	0.07	0.17	-	0.79	0.90	0.68	0.56	0.43	0.06	0.09	0.02	0.58	0.46	0.39
12	0.05	0.06	0.00	0.10	0.05	0.24	0.18	0.06	0.11	0.28	0.79	-	0.70	0.87	0.47	0.55	0.06	0.06	0.05	0.76	0.76	0.65
13	0.19	0.18	0.05	0.03	0.02	0.12	0.14	0.12	0.04	0.15	0.90	0.70	-	0.78	0.61	0.48	0.06	0.12	0.01	0.43	0.33	0.31
14	0.08	0.09	0.05	0.04	0.03	0.19	0.22	0.19	0.06	0.23	0.68	0.87	0.78	-	0.53	0.63	0.05	0.04	0.02	0.57	0.55	0.48
15	0.24	0.05	0.01	0.01	0.07	0.03	0.12	0.35	0.04	0.09	0.56	0.47	0.61	0.53	-	0.83	0.01	0.02	0.01	0.36	0.31	0.25
16	0.25	0.04	0.03	0.02	0.09	0.06	0.20	0.51	0.08	0.14	0.43	0.55	0.48	0.63	0.83	-	0.03	0.04	0.05	0.45	0.47	0.39
17	0.04	0.00	0.04	0.02	0.00	0.01	0.00	0.02	0.02	0.01	0.06	0.06	0.06	0.05	0.01	0.03	-	0.79	0.01	0.03	0.08	0.09
18	0.06	0.00	0.04	0.02	0.01	0.01	0.00	0.02	0.03	0.02	0.09	0.06	0.12	0.04	0.02	0.04	0.79	-	0.00	0.06	0.10	0.10
19	0.02	0.01	0.03	0.02	0.03	0.01	0.00	0.01	0.04	0.00	0.02	0.05	0.01	0.02	0.01	0.05	0.01	0.00	-	0.06	0.09	0.11
20	0.02	0.02	0.02	0.08	0.03	0.11	0.08	0.05	0.09	0.20	0.58	0.76	0.43	0.57	0.36	0.45	0.03	0.06	0.06	-	0.79	0.60
21	0.01	0.01	0.03	0.10	0.00	0.15	0.10	0.06	0.10	0.23	0.46	0.76	0.33	0.55	0.31	0.47	0.08	0.10	0.09	0.79	-	0.80
22	0.01	0.01	0.05	0.13	0.12	0.21	0.13	0.08	0.08	0.23	0.39	0.65	0.31	0.48	0.25	0.39	0.09	0.10	0.11	0.60	0.80	-

Variable 1, wellness score

Variable 2, number of days until the next match

Variable 3, age (years)

Variable 4, number of matches played in the prior season

Variable 5, session availability across pre-season (%)

Variable 6, total distance across pre-season (km)

Variable 7, high-speed running distance across pre-season (km)

Variable 8, sprint distance across pre-season (km)

Variable 9, isometric mid-thigh pull peak force (N/kg)

Variable 10, countermovement jump peak power (W/kg)

Variable 11, 6-day EWMA of total distance (m)

Variable 12, 28-day EWMA of total distance (m)

Variable 13, 6-day EWMA of HSR distance (m)

Variable 14, 28-day EWMA of HSR distance (m)

Variable 15, 6-day EWMA of sprint distance (m)

Variable 16, 28-day EWMA of sprint distance (m)

Variable 17, 6:28 day ratio of total distance

Variable 18, 6:28 day ratio of HSR distance

Variable 19, 6:28 day ratio of sprint distance

Variable 20, session availability in the prior 6 days (%)

Variable 21, session availability in the prior 28 days (%)

Variable 22, session availability in the prior 84 days (%)

EWMA, exponentially weighted moving average

HSR, high-speed running

Supplemental Digital Content 5. The coefficient and 95% confidence intervals (95% CIs) for all input variables, extracted from Full Model.

Variable	Coefficient	95% CIs
28-day EWMA of HSR distance (m)	-0.00298	-0.0037 to -0.00227
6-day EWMA of HSR distance (m)	0.00281	0.00247 to 0.00315
6-day EWMA of sprint distance (m)	-0.00255	-0.00651 to 0.00141
6:28 day ratio of HSR distance	-0.00354	-0.03787 to 0.0308
6:28 day ratio of sprint distance	0.00001	-0.00131 to 0.00134
6:28 day ratio of total distance	0.00966	-0.0177 to 0.03702
Age (years)	-0.07809	-0.25646 to 0.10028
Countermovement jump peak power (W/kg)	-0.01752	-0.13803 to 0.10299
Isometric mid-thigh pull peak force (N/kg)	0.03356	-0.09579 to 0.16292
Number of days until the next match	0.04231	0.0222 to 0.06242
Number of matches played in the prior season	0.01142	-0.07766 to 0.1005
Session availability across pre-season (%)	0.00832	-0.02989 to 0.04653
Session availability in the prior 6 days (%)	0.00401	0.00193 to 0.00608
Session availability in the prior 84 days (%)	0.00383	-0.00014 to 0.0078
Sprint distance across pre-season (km)	0.55167	0.01612 to 1.08722
Total distance across pre-season (km)	-0.01470	-0.03445 to 0.00505

EWMA, exponentially weighted moving average

HSR, high-speed running