Open access Original research

BMJ Open Sport & Exercise Medicine

Not straightforward: modelling non-linearity in training load and injury research

Lena Kristin Bache-Mathiesen ,[1] Thor Einar Andersen,[1] Torstein Dalen-Lorentsen ,[1] Benjamin Clarsen,[1,2] Morten Wang Fagerland[1,3]

To cite: Bache-Mathiesen LK, Andersen TE, Dalen-Lorentsen T, *et al*. Not straightforward: modelling non-linearity in training load and injury research. *BMJ Open Sport & Exercise Medicine* 2021;**7**:e001119. doi:10.1136/bmjsem-2021-001119

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjsem-2021-001119).

Accepted 16 July 2021

Check for updates

[1]Department of Sports Medicine, Oslo Sports Trauma Research Centre, Norwegian School of Sports Sciences, Oslo, Norway
[2]Centre for Disease Burden, Norwegian Institute of Public Health, Bergen, Norway
[3]Research Support Services, Oslo Centre for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway

Correspondence to
Lena Kristin Bache-Mathiesen; l.k.bache-mathiesen@nih.no

ABSTRACT

Objectives To determine whether the relationship between training load and injury risk is non-linear and investigate ways of handling non-linearity.

Methods We analysed daily training load and injury data from three cohorts: Norwegian elite U-19 football (n=81, 55% male, mean age 17 years (SD 1)), Norwegian Premier League football (n=36, 100% male, mean age 26 years (SD 4)) and elite youth handball (n=205, 36% male, mean age 17 years (SD 1)). The relationship between session rating of perceived exertion (sRPE) and probability of injury was estimated with restricted cubic splines in mixed-effects logistic regression models. Simulations were carried out to compare the ability of seven methods to model non-linear relationships, using visualisations, root-mean-squared error and coverage of prediction intervals as performance metrics.

Results No relationships were identified in the football cohorts; however, a J-shaped relationship was found between sRPE and the probability of injury on the same day for elite youth handball players (p<0.001). In the simulations, the only methods capable of non-linear modelling relationships were the quadratic model, fractional polynomials and restricted cubic splines.

Conclusion The relationship between training load and injury risk should be assumed to be non-linear. Future research should apply appropriate methods to account for non-linearity, such as fractional polynomials or restricted cubic splines. We propose a guide for which method(s) to use in a range of different situations.

Key messages

What is already known?

- ► Hypotheses suggest that the relationship between training load and injury risk is non-linear.
- ► Methods used in previous training load and injury research often assume linearity.
- ► Categorisation has been proven a suboptimal alternative for handling non-linearity.

What are the new findings?

- ► A non-linear relationship (p<0.001) between session rating of perceived exertion and the probability of injury in elite youth handball players would not have been discovered if linearity had been assumed (p=0.24).
- ► Acceptable Brier scores and C-statistics from a linear model do not mean that the relationship is linear.
- ► Categorising training load by quartiles could not model a linear relationship under skewed data conditions.
- ► Fractional polynomials and restricted cubic splines were the only methods capable of exploring non-linear shapes.

How might it impact clinical practice?

- ► Clinical researchers will have the tools available to perform causal and predictive research on training load and injury risk more accurately.
- ► More consistent methodology between training load and injury risk studies will improve comparability, reproducibility and facilitate meta-analyses.

INTRODUCTION

Injuries can hamper athlete and team performance in a variety of sporting disciplines.[1] Overuse injuries, in particular, are considered preventable, and in the last decade, researchers have investigated how training load affects injury risk in different football codes and other sports.[2] Results have been conflicting; some studies have found an increased risk with increased training loads, some have found that lower loads increase injury risk and some have found no association at all.[3 4] Hence, the relationship between training load and injury remains uncertain.

In 2013, Gamble theorised a U-shaped relationship between training load and injury risk. Too little and too much load increases risk,[5] with the middle section of the spectrum representing the lowest risk point. This hypothesis was revisited in 2016 by Blanch and Gabbett[6] who, based on three training load-injury datasets in different sports, postulated a workload–injury relationship that closely resembled a J-shaped curve; however, the statistical methodology in that paper has been questioned.[7] Gabbett[8] theorised a non-linear relationship between training load and injury risk with the rationale that training load may increase the risk of injury and build beneficial physiological adaptations such as aerobic capacity and strength, factors associated with decreased injury risk. The hypotheses of both Gamble and Gabbett suggest a non-linear relationship between different measures of training load and injury risk, prompting recent calls for better handling of non-linearity in the field.[9 10]

Despite these hypotheses and calls, methods that assume a linear relationship between training load and injury risk, such as Pearson correlations and logistic regression, are commonly used in the field.[11] If the training load and injury relationship is non-linear, such methods are expected to produce conflicting, irreproducible—and sometimes simply wrong—results. Nevertheless, no study has so far determined alternative methods for handling non-linearity.

The ideal method to handle non-linearity should be able to: (1) explore non-linear shapes and thus may confirm or reject previously outlined hypotheses; (2) model the non-linear relationship accurately; and (3) offer interpretable results.

The overall aim of this study was to identify the best methods for handling non-linearity in training load and injury research. First, we ascertained the relationship in three sports populations to reveal any potential evidence of non-linearity, to illustrate the problems and to present solutions. Second, we compared different methods in their ability to explore and accurately model potential non-linear shapes. Finally, we used the comparisons to develop a guide for which method(s) to use in different situations.

MATERIALS AND METHODS

Participants

We obtained training load and injury data collected from three cohorts: Norwegian elite U-19 football players (n=81, 55% male, mean age: 17 years, SD: 1 year),[12] one male football team from the Norwegian Premier League (n=36, mean age: 26 years (SD: 4))[13] and elite youth handball players recruited from Norwegian sports high schools (n=205, 36% male, mean age: 17 years (SD: 1)).[14] These cohorts were followed for 104, 323 and 237 days, respectively, during the competitive season.

All participants provided informed consent. Ethical principles were followed in accordance with the Declaration of Helsinki.

Training load definition

In all three cohorts, players reported the number of training sessions and matches daily. They also reported the duration of each activity and their rating of perceived exertion (RPE)[15] on the modified Borg CR10 scale.[16] To derive the session RPE (sRPE),[16] we multiplied the RPE by the activity duration in minutes.

Missing sRPE values are reported in online supplemental table S1 and were 24% for elite U-19 football, 41% for Premier League football and 64% for elite youth handball. The missing values were imputed using

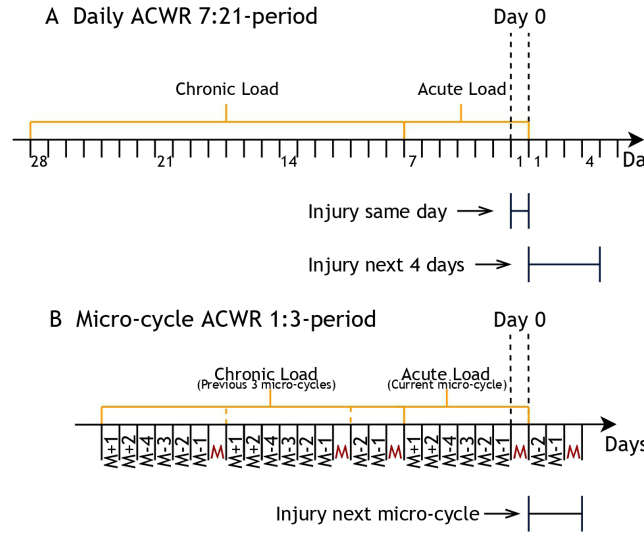

Figure 1 Illustration of time periods for calculating (A) daily ACWR 7:21-period and (B) micro-cycle ACWR 1:3-period. The first day that ACWR is calculated from is denoted day 0. The space between two tick marks represent 1 day (24 hours). For B, a microcycle period consists of all activity before a new match (M). That is, recovery days after the previous match as well as the training days before the next match. Days denoted with negative numbers are training days before the next match (M-1: being the day before the match; M-2: 2 days before a match and so on). Days with positive numbers are recovery and training days after a match (M+1: being the day after a match, M+2: 2 days after a match). The number of days between matches varies by the match schedule. How a team plan their training and recovery activities varies and is dependent on the teams' philosophy. For A, injury on the same day is defined as an injury on day 0, and future injury is defined as an injury occurring during the next 4 days excluding day 0. For B, future injury was defined as an injury occurring during the next microcycle excluding day 0. ACWR, acute:chronic workload ratio.

multiple imputation (online supplemental figure S1), a method that also performs well in cases of high amounts of missing (80%),[17] and the imputed values were deemed valid (online supplemental figure S2).

All load measures were based on players' daily ratings of perceived exertion (sRPE). We calculated an acute:chronic workload ratio (ACWR) in two different ways:

Daily ACWR 7:21-period

The mean sRPE across 7 days divided by the exponentially weighted moving average (EWMA) of the previous 21 days, uncoupled (figure 1).[18] The calculation was performed on a sliding window moving 1 day at a time from and including the 28th day.[19] The last day in the acute load is considered day 0 (figure 1).

Microcycle ACWR 1:3-period

The mean sRPE for each microcycle divided by the EWMA of the previous three microcycles uncoupled (figure 1). A microcycle was defined as all recovery days after the previous match and the training days before the next match. The next microcycle started on the first training day after the match and so on. For an illustration of a microcycle, see figure 1. The ACWR calculation was performed in the same manner as described for daily ACWR, on a sliding window moving one microcycle at a time from and including the fourth microcycle. The last day of the fourth microcycle was considered day 0 (figure 1).

When computing a ratio, one assumes that there is no relationship between the ratio and the denominator after controlling for the denominator; a ratio is only effective when the relationship between the numerator and the denominator is a straight line that intersects the origin.[20] Tests of this assumption are reported in online supplemental figure S3.

Injury definition

The same online questionnaire was used to collect daily health status and training information from all three sports cohorts. The elite U-19 football data and elite youth handball data were collected via the Briteback AB online survey platform, while the Norwegian Premier League football data were collected with Athlete Monitoring, Moncton. The players daily reported whether they had experienced 'no health problem', 'a new health problem' or 'an exacerbation of an existing health problem'. In the youth elite handball study, if players reported any new health problems, they were immediately prompted to specify whether it was an injury or illness in the questionnaire. In the football studies, if players reported any new health problems, a clinician contacted them by telephone the following day for a structured interview and classified the health problem as an injury or illness with the UEFA guidelines.[21] Players were asked to report all physical complaints, irrespective of their consequences on sports participation or the need to seek medical attention.[22]

Statistical analyses

To estimate the relationship between training load and injury risk, mixed effects logistic regression was used.[11 23]

We considered two outcomes: (1) occurrence of an injury on the same day as the observed training load (day 0) and (2) occurrence of injury in the future, where the current observation day (day 0) was not included. For unmodified training load values and daily ACWR 7:21-period, the future injury was defined as an injury occurring during the next 4 days excluding day 0. For microcycle ACWR 1:3-period, the future injury was any injury occurring during the next microcycle excluding day 0 (see figure 1 for an illustration of injury time periods and online supplemental table S2 for a list of the different models).

We adjusted for player age in all analyses. In addition, we adjusted for sex in the U-19 elite football and the elite youth handball models. In all models, the relationship between sRPE and injury risk was modelled with restricted cubic splines (RCSs).[24] The models were repeated without splines to simulate the relationship we would have discovered if we had assumed linearity. When using RCS, the estimated regression coefficients do not have a clinically meaningful interpretation, and only their p values are numerically interpretable.[24] The main result is, therefore, a visualisation of the model predictions (with uncertainty) to determine the shape of the relationship between training load and injury risk.

More details about data preparation and calculations are available in a supplementary file in .pdf format (online supplemental file 2). Our analyses served to illustrate whether there is any evidence for non-linearity in training load and injury research and should not be interpreted as causal inference.

Simulations

In addition to analysing real data, we performed (stochastic) simulations to compare different methods for ascertaining non-linear and linear relationships between training load and injury risk. The simulations were based on the elite U-19 football dataset since it had the least missing data (24%). The methodology here is focused on a causal research setting; however, the methods may also be applied in predictive research.[25] A detailed description of the simulation process and equations, as well as justifications for our methodological choices, is available as supplementary material (online supplemental file 2).

Two datasets were created. The first kept the original 8495 sRPE and 6308 ACWR values. In the second, sRPE and ACWR were sampled with replacement to generate 22 500 training load values.

Artificial injuries were simulated under different assumed scenarios for the relationship between training load and injury risk:

1. A U shape.
2. A J shape.
3. A linear shape.

A U shape between training load and injury risk indicates that the injury risk at lower levels of training load is equal to the injury risk at higher levels of training load. In contrast, moderate levels of training load have the lowest risk. In a J shape, moderate levels of training load have the lowest injury risk, followed by low levels of training load having intermediate risk. Finally, high levels of training load have the highest injury risk. For the U and linear relationship shapes, the simulated probability of an injury was based on the sRPE, while for the J shape, it was based on the ACWR. Any reference to the 'true' probability refers to the simulated probability we have created for a given scenario and which we aim to model.

We used mixed effects logistic regression models to estimate the relationship between training load and predefined injury risk, and we compared seven different methods to model the relationship:

- Linear model.
- Categorising by quartiles (data driven).
- Categorising by subjective cut-offs (subjective).
- Quadratic model.

- Fractional polynomials.
- RCSs with automated knots (data driven).
- RCSs with subjectively placed knots (subjective).

The root-mean-squared error (RMSE), coverage of prediction intervals, Brier score for model fit and C-statistics for predictive ability were calculated as performance measures. RMSE is a combined measure of accuracy and precision, where the lower the RMSE, the better the method. RMSE is only interpretable by comparing values in the same analysis – the values are meaningless in isolation.[26]

In summary, the four steps of the simulations were:

1. Sample training load values from the elite U-19 football data.
2. Simulate injuries with three different shapes for the relationship between injury risk and training load.
3. Fit seven different models with injury as the outcome and training load as the explanatory variable.
4. Calculate performance measures.

Steps 1–4 were repeated 1900 times.

For the U-shaped relationship, predicted values were visualised alongside the predefined shape to determine each method's ability to capture the true relationship. RMSE was also visually compared for the non-linear shapes.

All statistical analyses and simulations were performed using R V.4.0.2.[27] A GitHub repository is available with R code and data files.[28]

RESULTS

Evidence of non-linearity in training load and injury risk relationship research

A strong J-shaped relationship was found between sRPE and the probability of injury on the same day for elite youth handball players (p<0.001, figure 2A, online supplemental table S3). The linear model did not find this relationship (OR=1.0, 95% CI 0.99 to 1.00, p=0.24, figure 2B, online supplemental table S4). Additionally, for the handball cohort, an uncertain ∩-shaped relationship was present between sRPE and probability of injury in the next 4 days (p=0.06, figure 2B). These results also conflicted with the linear model showing no relationship (OR=1.0, 95% CI 0.99 to 1.00, p=0.35, figure 2B). For microcycle ACWR, the assumption that the relationship between the numerator and the denominator is a straight line intersecting the origin was supported, while for daily ACWR, the assumption was violated (online supplemental figure S3). No other relationships had significant p values or practically notable effect sizes (online supplemental table S3, figure S5 and S6).

Simulations

The quadratic model, fractional polynomials (FPs) and RCSs with subjectively placed knots were the only methods capable of modelling the non-linear U-shaped relationship (figure 3). FPs and RCS with subjectively placed knots (RCS subjectively) had the lowest RMSE and were, therefore, the best methods for the U shape (figure 4A).

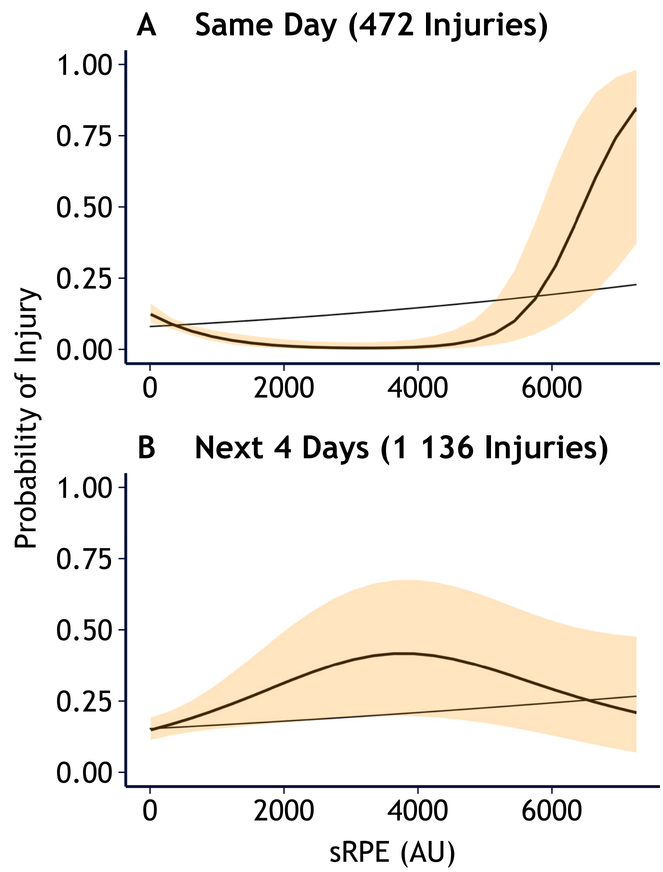

Figure 2 Probability of injury in elite youth handball on (A) the sameday and (B) the next 4 days, for each level of session rating of perceived exertion (sRPE) measured in arbitrary units (AU), as predicted by mixed effects logistic regression models with restricted cubic splines. The predictions pertain to a 17-year-old female. The yellow area represents 95% cluster-robust CIs around predicted values. The straight line shows the same predictions from an equivalent model without splines (ie, assuming linearity). For figure part B, modelling the response of injury in the next 4 days, multiple injuries on the same day were considered one event and an injury event would pertain to four load values and are therefore included four times.

The linear model had—by far—the highest RMSE and the data-driven RCS the second highest (figure 4A). In contrast, RCS (subjectively) had among the highest RMSE (figure 4B) regarding the J-shaped relationship. For the J shape, FPs and the quadratic model were the best methods (figure 4B). FPs had second-to-lowest RMSE for non-linear relationships (figure 4) and consistently had the best coverage (table 1).

All methods had a similar degree of error, predictive ability and model fit for the linear relationship (table 1).

The categorisation methods had the lowest coverage for the U and linear shapes, and categorising by quartiles had particularly poor coverage for the linear shape (25% vs >99% for other methods, table 1). For the J shape, the linear model performed worse than categorisation with 55% (vs 79% and 89%) for n=6308 (table 1). Predictions from the linear model could not form the U

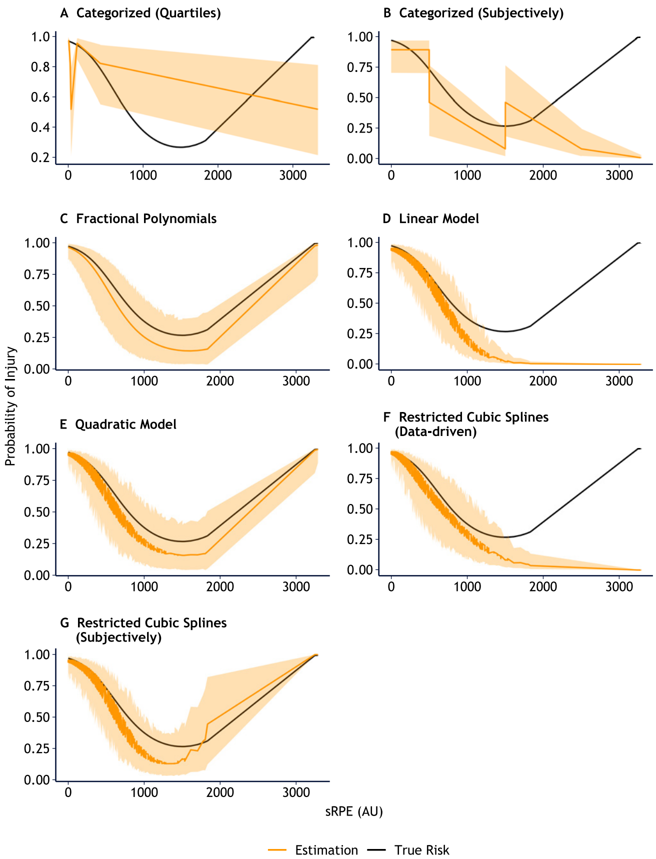

Figure 3 Probability of injury for each level of session rating of perceived exertion (sRPE) as predicted by seven different methods of modelling load. The yellow line represents the ability of the method to capture the U-shaped relationship (shown by the black line). The yellow area corresponds to the prediction interval. The predictions are based on 8494 sRPE values sampled from a highly skewed distribution in a Norwegian elite U-19 football dataset.

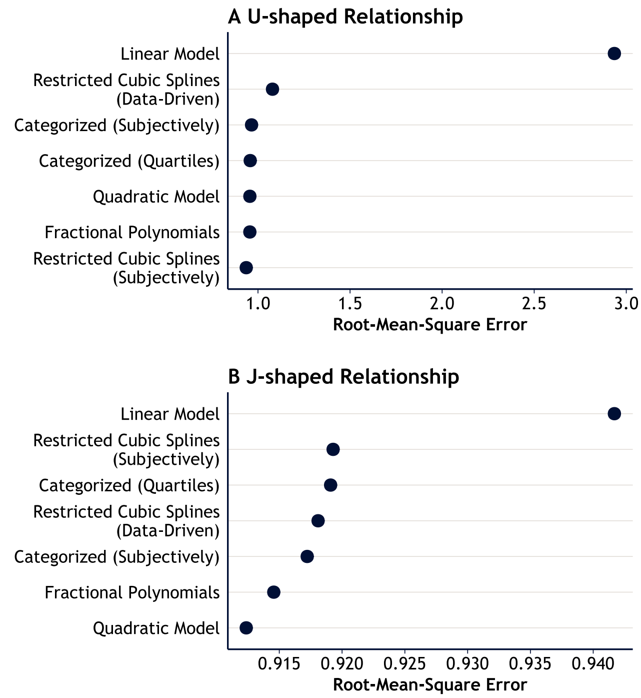

Figure 4 The mean root-mean-squared error (RMSE) of 1900 permutations for seven different methods modelling a non-linear (A) U-shaped relationship between session rating of perceived exertion (sRPE) and probability of injury, and (B) J-shaped relationship between acute:chronic workload ratio (ACWR) and probability of injury. The methods are arranged from top-to-bottom by the method with highest RMSE (most error) to the method with lowest RMSE. Thus, the best methods (those with lowest RMSE) are arranged towards the bottom. For figure part A, fractional polynomials and restricted cubic splines (subjectively) were the best methods, while for figure part B, fractional polynomials and the quadratic model were the best methods. The calculations are based on a Norwegian elite U-19 football dataset with 8494 sRPE values for (A) U shape and 6308 ACWR values for (B) J shape. RMSE cannot be compared between the two shapes, only within each shape.[26]

shape (figure 3) and had the highest degree of error for both non-linear shapes (highest RMSE; table 1, figure 4) but showed high predictive ability for the U shape (C-statistic >0.8) and moderate to poor predictive ability of the J shape (C-statistic=0.77 for n=6308, C-statistic=0.62 for n=22 500) in line with the other methods (table 1).

The differences in evaluation metrics between the two different sample sizes, n=22 500 and n=8494 for sRPE, and n=22 500 and n=6308 for ACWR, were negligible (table 1). Model fit determined by Brier score also failed to notably differentiate methods (table 1).

DISCUSSION

This is the first study exploring the potential for non-linearity in the relationship between training load and injury risk for football and handball. We found a J-shaped relationship between training load measured as the sRPE and probability of an injury on the same day in an elite youth handball cohort (figure 2A).

We also found that three methods were able to model the non-linear relationships between training load and injury explored in this paper: the quadratic model, FPs and RCSs, which managed to accurately recreate all simulated risk shapes (figure 4).

Evidence of non-linearity in training load and injury risk relationship research

All modelled relationships between training load and injury risk were either flat (no relationship) or non-linear. The results showed that the strength and direction of the relationship varied between training load—and injury—definitions in the handball population, while no relationships were found in the two football populations.

If we had assumed linearity and modelled the data accordingly, we would not have discovered these relationships. More grievously, we would have concluded there was no relationship between training load and injury risk for elite youth handball players for injury on the same day (linear model, p=0.24, type II error), when

Table 1 A comparison of mean root-mean-squared error, Brier score, C-statistic and coverage of prediction intervals for 1900 permutations of modelling the relationship between training load and risk of injury in seven different ways, with predetermined relationship shapes

Relationship	Sample size	Method	RMSE	Brier score	C-statistic	Coverage (%)
U shape	22 500	Linear model	2.344	0.097	0.827	100.000
		Categorised (quartiles)	0.995	0.101	0.809	99.678
		Categorised (subjectively)	0.996	0.102	0.758	94.600
		Quadratic model	0.993	0.097	0.826	100.000
		Fractional polynomials	0.994	0.096	0.829	100.000
		Restricted cubic splines (data driven)	1.065	0.097	0.826	100.000
		Restricted cubic splines (subjectively)	0.981	0.097	0.827	100.000
	8494	Linear model	2.935	0.093	0.851	98.048
		Categorised (quartiles)	0.958	0.096	0.838	98.769
		Categorised (subjectively)	0.965	0.098	0.809	84.600
		Quadratic model	0.956	0.092	0.850	98.937
		Fractional polynomials	0.956	0.092	0.852	98.942
		Restricted cubic splines (data driven)	1.079	0.092	0.849	98.686
		Restricted cubic splines (subjectively)	0.936	0.092	0.851	98.687
J shape	22 500	Linear model	1.044	0.063	0.618	77.694
		Categorised (quartiles)	0.993	0.064	0.689	88.652
		Categorised (subjectively)	0.993	0.063	0.690	96.404
		Quadratic model	0.984	0.061	0.732	99.997
		Fractional polynomials	0.986	0.061	0.740	100.000
		Restricted cubic splines (data driven)	0.992	0.061	0.735	99.999
		Restricted cubic splines (subjectively)	0.993	0.061	0.721	99.869
	6308	Linear model	0.942	0.060	0.774	54.493
		Categorised (quartiles)	0.919	0.060	0.791	79.120
		Categorised (subjectively)	0.917	0.059	0.795	89.393
		Quadratic model	0.912	0.057	0.817	93.272
		Fractional polynomials	0.915	0.057	0.821	95.517
		Restricted cubic splines (data driven)	0.918	0.057	0.818	94.281
		Restricted cubic splines (subjectively)	0.919	0.057	0.812	89.959
Linear	22 500	Linear model	0.999	0.239	0.591	100.000
		Categorised (quartiles)	0.999	0.240	0.588	25.000
		Categorised (subjectively)	0.999	0.241	0.579	99.995
		Quadratic model	0.999	0.239	0.591	99.999
		Fractional polynomials	0.999	0.239	0.592	100.000
		Restricted cubic splines (data driven)	0.999	0.239	0.591	100.000
		Restricted cubic splines (subjectively)	0.999	0.239	0.591	99.997
	8494	Linear model	0.991	0.228	0.655	99.795
		Categorised (quartiles)	0.991	0.228	0.653	24.957
		Categorised (subjectively)	0.991	0.229	0.649	99.678
		Quadratic model	0.991	0.228	0.656	99.786
		Fractional polynomials	0.991	0.228	0.656	99.788
		Restricted cubic splines (data driven)	0.991	0.228	0.656	99.789
		Restricted cubic splines (subjectively)	0.991	0.228	0.656	99.791

RMSE, root-mean-squared error.

it was, in fact, a strong U-shaped parabola (RCS model, $p<0.001$, figure 2A). This may happen when a relationship is not only non-linear but non-monotonic. In monotonic relationships, the response variable Y (injury probability) moves only in one direction as X (training load) increases, while in non-monotonic relationships, Y sometimes increases and sometimes decreases when X increases.[9]

In 2013, Gamble[5] theorised a U-shaped relationship between training load and risk of injury. Data presented by Blanch and Gabbett[6] suggested a J-shaped relationship between ACWR and injury, although the methodology and interpretation of this finding have recently been questioned.[7] Here, we reproduced a J shape between sRPE and injury occurring on the same day for elite youth handballers but not for the relative training load described by the ACWR in the same cohort. In Lathlean *et al*,[29] a U shape was discovered between training load and the risk of future injury in an Australian football cohort. These findings might suggest that the training load and injury relationship is different for different sports and populations. Since non-linearity is possible in a training load and injury context, we recommend assuming the data have an unknown, non-linear relationship when conducting statistical analyses.

Methods for addressing non-linear relationships

As expected, standard logistic regression could not model the U and J shapes, as it assumes linearity. For the U shape, the RMSE was threefold higher for the linear model than all other models (RMSE=2.9 vs RMSE≈0.95, figure 4A), showing that violation of the linearity assumption causes major bias and can substantially alter conclusions based on the results. Misleadingly, the linear model had a great C-statistic score (>0.8) and comparable Brier scores. This happened because the sRPE values were highly skewed (online supplemental figure S4). Over 90% of the data points were congested in the left-hand side of the U shape (figure 3, online supplemental figure S4). The linear model, which only managed to model the left-hand side of the U shape, therefore predicted most of the values well, causing the impressive C-statistic. However, it could not predict the right-hand side of the U shape at all and therefore had high RMSE. Consequently, a researcher who measures model fit by predictive ability alone may be falsely assured that the linearity assumption holds true.

Categorisation has previously been explored thoroughly in Carey *et al*[30] and proven a poor method for modelling non-linear relationships. The results were reproduced in our study using a football population, where the RMSE and coverage for categorisation were consistently outperformed by other methods (table 1). In addition, our results showed that categorising by quartiles was suboptimal for modelling non-linear relationships and also suboptimal when the relationship between training load and injury risk was linear (coverage of 25% vs >99% for all other methods).

Recently, some studies have added a quadratic term to the training load and injury model to test for linearity: if the term was non-significant, it was discarded for a linear model; if significant, they categorised the training load variable to handle non-linearity.[31–33] If the quadratic term is significant, the researchers correctly choose other options over a linear model. However, the quadratic term only tests for a parabolic shape—not non-linearity in general. A significant quadratic term does not mean the relationship is quadratic (parabolic). It means that a quadratic shape fits better than a linear shape. If the quadratic term is not significant, it does not necessarily mean the underlying relationship is linear, either, only that a quadratic shape fits poorly. Furthermore, testing non-linearity with a quadratic term has been shown to inflate type I error rates by 50%.[34]

Blanch and Gabbett[6] and Carey *et al*[19] used quadratic regression assuming a parabolic relationship between training load and injury risk. In our study, quadratic regression modelled the U-shaped risk profiles with low degrees of error (figures 3 and 4A) and had the best performance for the J-shaped relationship (figure 4B). This is expected, as the J shape was initially constructed from a quadratic model in Blanch and Gabbett.[6] Contrary to a real-life setting, however, we knew the risk profiles before analysing our data. Quadratic regression does not explore shapes but constrains the model to follow a specific pathway. We think it is only appropriate when strong evidence from previous studies support a parabolic relationship. We recommend assuming non-linearity of unknown shapes and using methods not to test for linearity but to explore and model non-linearity to discover the relationship. Based on our findings and previous research in other fields such as medical statistics,[35] FPs and RCSs appear to be the best methods for doing this.

FPs modelled all risk shapes accurately (figure 4, table 1). FP has recently been used in a training load and injury risk study.[29] This method requires minor subjective influence, and the results are intuitive, especially for users familiar with quadratic regression. Although it appears the superior choice at first glance, the method has a disadvantage: FPs are defined only for positive values, which means that an FP model is unable to model negative values and the value 0. In the context of training load and injury risk research, training load is (traditionally) never measured on a negative scale.[36] If it can be justified, adding a small constant (such as 0.001, or whatever is considered small in the context of the measuring scale) to all training load values can solve the problem with 0 and allow the use of FPs.

RCSs performance depended on how knot locations (the points where the polynomials that make up cubic splines are joined, see online supplemental file 2 for details) were chosen. In the data-driven method, where knots were automatically placed by the default setting, RCS failed to model the U-shaped scenario (figure 3). When knot position was chosen based on the range of

Box 1 Recommended methods to model non-linear relationships between training load and injury risk

To model non-linear relationships, either Fractional Polynomials (FP) or Restricted Cubic Splines (RCS) can be used.
Fractional polynomials are easier to interpret. We recommend FP under the following conditions:

- When the main objective is causal research, FP is preferred. When the training load measure does not include negative numbers or 0. This includes:
  - Studies that use the Acute-Chronic Workload Ratio or other metrics that cannot be the value 0 or a negative value.
  - Studies that model the relationship between training load and injury risk on the same day, or other scenarios where the researchers may wish to remove the days where the athletes were not exposed to any training load from the dataset.
  - Studies that can justify applying a small constant (such as 0.001, or whatever is considered small in the context of the measuring scale) to all training load values.

We recommend restricted cubic splines under the following conditions:

- When the main objective is predictive research, RCS is preferred.
- When the training load measures must have the value 0. This includes studies that wish to capture a change in the effect, regardless how small, going from no training load at all to any amount of training load.
- When training load is included in the study merely to adjust for it as a potential confounder and is not the main variable of interest.

We do not recommend changing the study aims or the chosen measure to use FP, nor do we recommend using FP under certain conditions and RCS for other conditions in the same study.
A step-by-step guide to performing FP and RCS in R can be accessed on the primary author's GitHub page.[39 40]

the training load variable, RCS modelled the U accurately (figure 3). However, the results were the opposite for the J-shaped relationship where the data-driven method was among those of lowest error, and the subjectively located knots had the highest amount of error (figure 4B). The default placement algorithm was by quartiles, and in the highly skewed distribution of the sRPE values used in the U-shaped relationship (online supplemental figure S4), it caused the knots to be placed tightly together (figure 3). Therefore, it could not model the shape, while the subjective version was created with the range of the values in mind. The ACWR values used in the J shape had a Gaussian distribution (online supplemental figure S4), and using quartiles was a feasible choice. This shows the importance of careful model calibration using clinical knowledge and knowledge of the data.

RCS produces effect sizes that are difficult to use in a practical setting, and results can only be interpreted in the form of p values and visualisation (such as in figure 2). RCS is less ideal than FP in causal research. Still, its disadvantages are not as relevant in predictive research where interpretability is of minor concern.[25] We propose a guide for when FP is recommended and when RCS is recommended (box 1).

Limitations

A limitation of this study was the sample size, the number of injuries and consequently statistical power. Neither of the two football cohorts satisfied the recommendation of >200 injuries to detect a small to moderate effect.[37] The elite youth handball data, despite having a sufficient number of injuries, had high amounts of missing sRPE values (64%), and this may have caused selection bias. We emphasise that the exploration of non-linearity in these data were for illustrative purposes and not to show causal inference.

We used statistical methods commonly used and recommended in the field to demonstrate how non-linear relationships can be ascertained with existing methods. We were consequently limited in the choice of methods. The ACWR model is under debate, and the pros and cons of the method have been explored extensively in recent publications.[12 18 38] The purpose of this paper was not to provide additional insight into that discussion but rather to demonstrate how a continuous training load variable should be modelled to account for non-linearity. For this reason, we opted to use ACWR, as it is currently the most used training load method in the field of training load and injury risk research.[4]

CONCLUSION

Exploratory analyses showed evidence of a non-linear relationship between training load and risk of injury in a sports population. Researchers should assume that the relationship between training load and injury risk is non-linear and use appropriate methods that explore relationships rather than constrain them. Linear methods should only be used when the relationship is first proven to be linear. We promote FPs or RCSs to model non-linear relationships, depending on the scenario.

Twitter Lena Kristin Bache-Mathiesen @lena_kbm, Thor Einar Andersen @DocThorAndersen, Torstein Dalen-Lorentsen @torsteindalen, Benjamin Clarsen @benclarsen and Morten Wang Fagerland @FagerlandWang

Acknowledgements We would like to thank Christian Thue Bjørndal for access to the elite youth handball data. We would also like to thank Garth Theron for high-quality programming of the Norwegian Premier League football database. This research would not be made possible without the collaboration of coaches and athletes, and we would like to thank the participants who contributed.

Contributors LKB-M designed the study and performed statistical analyses in collaboration with and under supervision from MWF and TEA. TDL constructed the novel idea of using microcycles instead of calendar weeks. All authors contributed with notable critical appraisal of the text and approved the final version.

Funding The Oslo Sports Trauma Research Centre provided all funding for performing this study.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval Study protocol for all three studies were approved by the Norwegian Centre for Research Data: Norwegian elite U-19 football (5487), Norwegian Premier League football (722773) and Norwegian elite youth handball (407930). They were also approved by the Ethical Review Board of the Norwegian School of Sport Sciences. The Norwegian elite U-19 football study was also approved by the South-Eastern Norway Regional Committee for Medical and Health Research Ethics (2017/1015).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available on reasonable request. Data used for simulations are available in a public, open access repository (https://github.com/lenakba/load-injury-non-linearity-study). The Norwegian elite U-19 football data, Norwegian Premier League football data and Norwegian elite youth handball data are available on reasonable request. These are anonymised based on requirements of the Norwegian Data Protection Agency. The removal of background variables for the anonymisation renders the data unusable for any reproducibility purposes; the data are only available for the sake of transparency.

ORCID iDs
Lena Kristin Bache-Mathiesen http://orcid.org/0000-0002-9466-7003
Torstein Dalen-Lorentsen http://orcid.org/0000-0003-4062-7601

REFERENCES

1 Hägglund M, Waldén M, Magnusson H, *et al*. Injuries affect team performance negatively in professional football: an 11-year follow-up of the UEFA champions League injury study. *Br J Sports Med* 2013;47:738–42.
2 Griffin A, Kenny IC, Comyns TM. The association between the acute: chronic workload ratio and injury and its application in team sports: a systematic review. *Sports Medicine* 2019:1–20.
3 Eckard TG, Padua DA, Hearn DW, *et al*. The relationship between training load and injury in athletes: a systematic review. *Sports Medicine* 2018;48:1929–61.
4 Andrade R, Wik EH, Rebelo-Marques A, *et al*. Is the acute: chronic workload ratio (ACWR) associated with risk of Time-Loss injury in professional team sports? A systematic review of methodology, variables and injury risk in practical situations. *Sports Med* 2020;50:1–23.
5 Gamble P. Reducing injury in elite sport-is simply restricting workloads really the answer? *N Zeal J Sports Med* 2013;36.
6 Blanch P, Gabbett TJ. Has the athlete trained enough to return to play safely? The acute:chronic workload ratio permits clinicians to quantify a player's risk of subsequent injury. *Br J Sports Med* 2016;50:471–5.
7 Impellizzeri F, Wookcock S, McCall A. *The acute-chronic workload ratio-injury figure and its 'sweet spot'are flawed*, 2019.
8 Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter *and* harder? *Br J Sports Med* 2016;50:273–80.
9 Stern BD, Hegedus EJ, Lai Y-C. State dependence: does a prior injury predict a future injury? *Phys Ther Sport* 2021;49:8–14.
10 Kalkhoven JT, Watsford ML, Coutts AJ. Training load and injury: causal pathways and future directions. *Sports Med* 2021:1–14. doi:10.1007/s40279-020-01413-6
11 Windt J, Ardern CL, Gabbett TJ, *et al*. Getting the most out of intensive longitudinal data: a methodological review of workload–injury studies. *BMJ Open* 2018;8:e022626.
12 Dalen-Lorentsen T, Andersen TE, Bjørneboe J. A cherry tree ripe for picking: the relationship between the acute: chronic workload ratio and health problems. *J Orthop Sports Phys Ther* 2021;0:1–26.
13 Theron GF. *The use of data mining for predicting injuries in professional football players*. University of Oslo, 2020.
14 Bjørndal CT, Bache-Mathiesen LK, Gjesdal S, *et al*. An examination of training load, match activities, and health problems in Norwegian youth elite Handball players over one competitive season. *Front Sports Act Living* 2021;3:635103.
15 Borg G, Hassmén P, Lagerström M. Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *Eur J Appl Physiol Occup Physiol* 1987;56:679–85.
16 Foster C, Florhaug JA, Franklin J, *et al*. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15:10Y 115.
17 Madley-Dowd P, Hughes R, Tilling K, *et al*. The proportion of missing data should not be used to guide decisions on multiple imputation. *J Clin Epidemiol* 2019;110:63–73.
18 Wang C, Vargas JT, Stokes T, *et al*. Analyzing Activity and Injury: Lessons Learned from the Acute:Chronic Workload Ratio. *Sports Medicine*2020;50:1243–54.
19 Carey DL, Blanch P, Ong K-L, *et al*. Training loads and injury risk in Australian football-differing acute: chronic workload ratios influence match injury risk. *Br J Sports Med* 2017;51:1215–20.
20 Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. *Adv Physiol Educ* 2013;37:213–9.
21 Hägglund M*et al*. Methods for epidemiological study of injuries to professional football players: developing the UEFA model. *Br J Sports Med* 2005;39:340–6.
22 Fuller CW, Ekstrand J, Junge A, *et al*. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Scand J Med Sci Sports* 2006;16:83–92.
23 Nielsen Rasmus Østergaard, Shrier I, Casals M, *et al*. Statement on methods in sport injury research from the first methods matter meeting, Copenhagen, 2019. *J Orthop Sports Phys Ther* 2020;50:226–33.
24 Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med* 1989;8:551–61.
25 Shmueli G. To explain or to predict? *Statistical Science* 2010;25:289–310.
26 Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med* 2019;38:2074–102.
27 R Core Team. *R foundation for statistical computing, Vienna, Austria. 4.0.2*, 2013. http://www.R-project.org/
28 Bache-Mathiesen L. Modelling non-linearity in training load and injury research: R code Repository, 2021. Available: https://github.com/lenakba/load-injury-non-linearity-study [Accessed 12 Apr 2021].
29 Lathlean TJH, Gastin PB, Newstead SV, *et al*. Absolute and relative load and injury in elite junior Australian football players over 1 season. *Int J Sports Physiol Perform* 2019;1:1–9.
30 Carey DL, Crossley KM, Whiteley R, *et al*. Modeling training loads and injuries: the dangers of Discretization. *Med Sci Sports Exerc* 2018;50:2267–76.
31 Warren A, Williams S, McCaig S, *et al*. High acute:chronic workloads are associated with injury in England & Wales Cricket Board Development Programme fast bowlers. *J Sci Med Sport* 2018;21:40–5.
32 Sampson JA, Murray A, Williams S, *et al*. Injury risk-workload associations in NCAA American College football. *J Sci Med Sport* 2018;21:1215–20.
33 Ahmun R, McCaig S, Tallent J, *et al*. Association of daily workload, wellness, and injury and illness during Tours in international cricketers. *Int J Sports Physiol Perform* 2019;14:369–77.
34 Grambsch PM, O'Brien PC. The effects of transformations and preliminary tests for non-linearity in regression. *Stat Med* 1991;10:697–709.
35 Binder H, Sauerbrei W, Royston P. Comparison between splines and fractional polynomials for multivariable model building with continuous covariates: a simulation study with continuous response. *Stat Med* 2013;32:2262–77.
36 Bourdon PC, Cardinale M, Murray A, *et al*. Monitoring athlete training loads: consensus statement. *Int J Sports Physiol Perform* 2017;12:S2-161-S2-70
37 Bahr R, Holme I. Risk factors for sports injuries--a methodological approach. *Br J Sports Med* 2003;37:384–92.
38 Impellizzeri FM, Tenan MS, Kempton T, *et al*. Acute:Chronic Workload Ratio: Conceptual Issues and Fundamental Pitfalls. *Int J Sports Physiol Perform* 2020;15:907–13.
39 Bache-Mathiesen L. Performing Fractional Polynomials in R [PDF]. GitHub, 2021. Available: https://github.com/lenakba/load-injury-non-linearity-study/blob/main/guide_fractional_polynomials.pdf [Accessed 12 Apr 2021].
40 Bache-Mathiesen L. Performing Restricted Cubic Splines in R [PDF]. GitHub, 2021. Available: https://github.com/lenakba/load-injury-non-linearity-study/blob/main/guide_restricted_cubic_splines.pdf [Accessed 12 Apr 2021].

SUPPLEMENTARY RESULTS

Tables

Table S1. Data quality comparison of sports cohorts: the Norwegian elite U-19 football data (55% male, age; mean ± standard deviation (SD) = 17 ± 1 years), Norwegian Premier League football data (all male, age 26 ± 4 years) and elite youth handball data (36% male, age 17 ± 0.9 years).

		Football U-19	Football Elite	Handball
Sample Size	Number of athletes	81	36	205
	Number of sRPE values before imputation	6 424	6 061	17 268
	Number of sRPE values after imputation	8 495	10 232	47 651
	Number of injuries	81	38	472
	Number of injuries per athlete, mean (SD)	1 (1.2)	1 (1.5)	2.3 (2.9)
Missing data	Missing load values, n (%)	2 071 (24%)	4 171 (41%)	30 383 (64%)
	Missing load values per athlete, mean (SD)	26 (32)	116 (62)	148 (71)
Timelines	Mean (SD) answering time, days	0.3 (0.7)	0.01 (0.2)	0.7 (1.6)
	Percentage of forms answered the same day	72%	99%	53%
	Max answering time, days	9	4	119

Abbreviations: Football Elite, Norwegian Premier League; sRPE, session Rating of Perceived Exertion

Bache-Mathiesen LK, *et al. BMJ Open Sp Ex Med* 2021; 7:e001119. doi: 10.1136/bmjsem-2021-001119

Table S2. Overview of injury definition and models run on each sport population, with the number of load values and the number of injuries used in each model.

Population	Injury Definition[1]	Load Definition[2]	Load Values (n)[3]	Injuries (n)[3]
Football U-19 (n = 81)	Same day	sRPE	8495	81
		Daily ACWR 7:21-period	6308	43
	Next 4 days	sRPE	8495	210
		Daily ACWR 7:21-period	6308	129
	Next micro-cycle	Micro-cycle ACWR 1:3-period	793	26
Football Elite (n = 36)	Same day	sRPE	10 232	38
		Daily ACWR 7:21-period	9 260	32
	Next 4 days	sRPE	10 232	44
		Daily ACWR 7:21-period	9 260	34
	Next micro-cycle	Micro-cycle ACWR 1:3-period	553	26
Handball (n = 205)	Same day	sRPE	47 651	472
		Daily ACWR 7:21-period	42 116	320
	Next 4 days	sRPE	47 651	1 136
		Daily ACWR 7:21-period	42 116	714
	Next micro-cycle	Micro-cycle ACWR 1:3-period	1 897	242

Abbreviations: ACWR, Acute: Chronic Workload Ratio; Football Elite, Norwegian Premier League; sRPE = daily session Rating of Perceived Exertion; TL, Training Load.

[1]Same day was injury same day as the measured load value; Next 4 days was one or more injuries during the four days after the measured load value; Next micro-cycle was one or more injuries during the micro-cycle after the micro-cycle of the measured load values.

[2]Daily ACWR 7:21-period was the 7-day acute sRPE divided by previous 21-day chronic sRPE per day; Micro-cycle ACWR 1:3-period was the 1-micro-cycle acute sRPE divided by previous 3-micro-cycle chronic sRPE per micro-cycle. A micro-cycle was defined as all recovery days after the previous match as well as the training days before the next match.

[3]Due to aggregations, ACWR calculations and injury time-windows, the number of load values and injury events varied between models.

Table S3. Odds ratio with 95% confidence intervals, standard error, degrees of freedom and p-values from modelling the relationship between training load and injury risk using mixed effect models with restricted cubic splines.

Population	Load Definition[1]	Injury Definition[2]	Variable	OR[3]	CI 2.5%	CI 97.5%	SE	df	p
Football U-19	sRPE	Same day	Intercept	0.004	<0.001	0.939	2.808	941	0.047
			Load	1.000	0.997	1.003	0.002	3331	0.837
			Load'	1.001	0.997	1.004	0.002	3337	0.746
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.163	0.628	2.155	0.314	3286	0.631
			Age (Years)	1.088	0.800	1.479	0.157	921	0.592
		Next 4 days	Intercept	0.031	<0.001	13.614	3.094	273	0.262
			Load	1.001	1.000	1.002	0.001	4253	0.179
			Load'	0.999	0.998	1.001	0.001	3386	0.502
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.181	0.582	2.400	0.362	4727	0.645
			Age (Years)	0.973	0.689	1.373	0.175	261	0.874
	Daily ACWR 7:21-period	Same day	Intercept	0.002	<0.001	1.073	3.233	1088	0.053
			Load	0.778	0.313	1.936	0.465	1896	0.589
			Load'	2.970	0.586	15.057	0.827	1268	0.189
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.326	0.648	2.716	0.365	2592	0.440
			Age (Years)	1.118	0.785	1.594	0.181	1131	0.536
		Next 4 days	Intercept	<0.001	<0.001	25.567	6.179	104	0.148
			Load	4.285	1.241	14.793	0.631	498	0.021
			Load'	0.032	0.007	0.139	0.745	1565	<0.001
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.278	0.381	4.283	0.617	2328	0.691
			Age (Years)	1.160	0.594	2.266	0.338	110	0.661

Continues on the next page

Population	Load Definition[1]	Injury Definition[2]	Variable	OR[3]	CI 2.50 %	CI 97.50 %	SE	df	p
Football U-19	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	0.041	<0.001	40.086	3.502	396	0.362
			Load	0.210	0.012	3.536	1.439	562	0.278
			Load'	8.535	<0.001	6136037	6.857	356	0.755
			Load''	0.144	<0.001	3.52E+20	25.026	296	0.938
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.107	0.478	2.563	0.427	641	0.812
			Age (Years)	1.070	0.732	1.563	0.193	358	0.726
Football Elite	sRPE	Same day	Intercept	0.001	<0.001	0.011	1.437	4480	<0.001
			Load	1.000	0.995	1.005	0.003	4479	0.897
			Load'	1.001	0.994	1.008	0.004	4475	0.847
			Age (Years)	1.096	0.997	1.204	0.048	4480	0.056
		Next 4 days	Intercept	<0.001	<0.001	0.022	2.581	1593	0.001
			Load	0.998	0.994	1.003	0.002	168	0.501
			Load'	1.004	0.997	1.011	0.003	99	0.29
	Daily ACWR 7:21-period	Same day	Intercept	<0.001	<0.001	0.022	2.465	55	0.001
			Load	3.389	0.042	273.286	2.119	22	0.57
			Load'	0.337	0.004	31.613	2.201	24	0.626
			Age (Years)	1.104	0.994	1.226	0.053	3833	0.064
			Age (Years)	1.186	0.991	1.418	0.091	1662	0.062
		Next 4 days	Intercept	<0.001	<0.001	0.015	3.485	300	0.002
			Age (Years)	1.202	0.978	1.477	0.105	1349	0.081
			Load	6.731	0.116	390.17	1.948	20	0.339
			Load'	0.056	0.001	5.583	2.252	29	0.21
	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	<0.001	<0.001	0.136	2.841	62	0.009
			Age (Years)	1.113	1.016	1.219	0.046	476	0.021
			Load	7.523	0.030	1881.323	2.742	46	0.466
			Load'	0.340	0.005	22.344	2.113	112	0.610

Continues on the next page

Population	Load Definition[1]	Injury Definition[2]	Variable	OR[3]	CI 2.50 %	CI 97.50 %	SE	df	p
Handball	sRPE	Same day	Intercept	0.083	0.003	2.711	1.777	1632	0.162
			Load	0.999	0.998	0.999	<0.001	9445	<0.001
			Load'	1.002	1.001	1.003	0.001	2603	<0.001
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.112	0.780	1.586	0.181	11867	0.556
			Age (Years)	0.963	0.787	1.177	0.102	1740	0.711
		Next 4 days	Intercept	0.606	0.007	54.891	2.297	1270	0.827
			Load	1.000	1.000	1.001	<0.001	39	0.063
			Load'	0.999	0.999	1.000	<0.001	21	0.143
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.053	0.645	1.719	0.25	11521	0.837
			Age (Years)	0.87	0.67	1.129	0.133	1146	0.294
	Daily ACWR 7:21-period	Same day	Intercept	0.041	0.001	2.833	2.157	3372	0.140
			Load	0.743	0.362	1.523	0.366	1301	0.417
			Load'	1.648	0.687	3.952	0.445	394	0.262
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.127	0.729	1.741	0.222	8737	0.591
			Age (Years)	0.989	0.776	1.259	0.124	3357	0.926
		Next 4 days	Intercept	0.234	0.001	99.719	3.088	2022	0.638
			Load	2.006	1.006	4.002	0.348	98	0.048
			Load'	0.292	0.133	0.643	0.395	70	0.003
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.316	0.708	2.449	0.317	7490	0.385
			Age (Years)	0.886	0.624	1.257	0.179	1426	0.497

Continues on the next page

Population	**Load Definition[1]**	**Injury Definition[2]**	**Variable**	**OR[3]**	**CI 2.50 %**	**CI 97.50 %**	**SE**	**df**	**p**
Handball	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	0.165	0.003	9.425	2.062	1450	0.382
			Load	0.878	0.397	1.939	0.404	955	0.747
			Load'	1.335	0.599	2.976	0.408	969	0.479
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	0.908	0.596	1.384	0.215	1551	0.654
			Age (Years)	1.004	0.795	1.267	0.119	1313	0.976

Abbreviations: ACWR, Acute: Chronic Workload Ratio; CI, 95% Confidence Intervals; df, Degrees of Freedom; Football Elite, Norwegian Premier League; OR, Odds Ratio; SE, Standard Error; sRPE, daily session Rating of Perceived Exertion.

[1]Daily ACWR 7:21-period was the 7-day acute sRPE divided by previous 21-day chronic sRPE per day; Micro-cycle ACWR 1:3-period was the 1-micro-cycle acute sRPE divided by previous 3-micro-cycle chronic sRPE per micro-cycle. A micro-cycle was defined as all recovery days after the previous match as well as the training days before the next match.

[2]Same day was injury same day as the measured load value; Next 4 days was one or more injuries during the four days after the measured load value; Next micro-cycle was one or more injuries during the micro-cycle after the micro-cycle of the measured load values.

[3]As load was fitted with cubic splines, the effect-size, Odds Ratio, is uninterpretable for this parameter.

Bache-Mathiesen LK, *et al. BMJ Open Sp Ex Med* 2021; 7:e001119. doi: 10.1136/bmjsem-2021-001119

Table S4. Odds ratio with 95% confidence intervals, standard error, degrees of freedom and p-values from modelling the relationship between training load and injury risk using mixed effect logistic regression models which assume linearity.

Population	**Load Definition**[1]	**Injury Definition**[2]	**Variable**	**OR**	**CI 2.5%**	**CI 97.5%**	**SE**	**df**	**p**
Football U-19	sRPE	Same day	Intercept	0.003	<0.001	0.795	2.781	942	0.041
			Load	1.000	0.999	1.001	<0.001	3338	0.755
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.165	0.629	2.156	0.314	3287	0.627
			Age (Years)	1.088	0.800	1.478	0.156	922	0.591
		Next 4 days	Intercept	0.034	<0.001	15.046	3.099	263	0.275
			Load	1.000	1.000	1.001	<0.001	3015	0.067
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.186	0.585	2.406	0.361	4759	0.636
			Age (Years)	0.969	0.686	1.369	0.175	256	0.86
	Daily ACWR 7:21-period	Same day	Intercept	0.001	<0.001	0.405	3.181	1341	0.025
			Load	1.346	0.859	2.107	0.228	492	0.194
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.329	0.645	2.737	0.368	2601	0.440
			Age (Years)	1.139	0.798	1.625	0.181	1289	0.473
		Next 4 days	Intercept	0.003	<0.001	98.996	5.313	108	0.266
			Load	1.312	0.633	2.722	0.372	1744	0.465
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.408	0.478	4.147	0.551	2514	0.535
			Age (Years)	1.052	0.583	1.899	0.298	108	0.865

Continues on the next page

Population	Load Definition[1]	Injury Definition[2]	Variable	OR	CI 2.5%	CI 97.5%	SE	df	p
Football U-19	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	0.013	<0.001	10.059	3.396	331	0.199
			Load	0.850	0.324	2.232	0.492	534	0.741
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.088	0.475	2.492	0.422	645	0.842
			Age (Years)	1.079	0.738	1.577	0.193	299	0.694
Football Elite	sRPE	Same day	Intercept	0.001	<0.001	0.008	1.344	4481	<0.001
			Load	1.000	0.999	1.002	0.001	4481	0.867
			Age (Years)	1.096	0.998	1.204	0.048	4481	0.055
		Next 4 days	Intercept	<0.001	<0.001	0.014	2.572	1412	<0.001
			Load	1.001	0.998	1.003	0.001	15	0.484
			Age (Years)	1.189	0.994	1.422	0.091	1662	0.058
	Daily ACWR 7:21-period	Same day	Intercept	<0.001	<0.001	0.008	1.555	3576	<0.001
			Load	1.255	0.459	3.427	0.512	552	0.658
			Age (Years)	1.102	0.994	1.222	0.052	3847	0.064
		Next 4 days	Intercept	<0.001	<0.001	0.042	2.952	1288	0.002
			Load	0.739	0.253	2.165	0.539	70	0.577
			Age (Years)	1.189	0.974	1.452	0.102	1356	0.089
	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	0.001	<0.001	0.036	1.665	186	<0.001
			Load	2.183	0.350	13.625	0.910	47	0.396
			Age (Years)	1.115	1.018	1.221	0.046	476	0.019
Handball	sRPE	Same day	Intercept	0.063	0.002	2.082	1.782	1673	0.121
			Load	1.000	1.000	1.000	0.000	7341	0.240
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.124	0.788	1.604	0.181	11869	0.519
			Age (Years)	0.953	0.779	1.166	0.103	1739	0.638

Continues on the next page

Population	Load Definition[1]	Injury Definition[2]	Variable	OR	CI 2.5%	CI 97.5%	SE	df	p
Handball	sRPE	Next 4 days	Intercept	0.603	0.006	56.793	2.317	1367	0.827
			Load	1.000	1.000	1.000	<0.001	64	0.348
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.045	0.641	1.705	0.250	11524	0.859
			Age (Years)	0.873	0.671	1.135	0.134	1420	0.310
	Daily ACWR 7:21-period	Same day	Intercept	0.030	<0.001	2.084	2.159	2373	0.105
			Load	1.106	0.846	1.445	0.136	204	0.459
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.129	0.730	1.748	0.223	8738	0.585
			Age (Years)	0.988	0.775	1.260	0.124	3373	0.921
		Next 4 days	Intercept	0.535	0.001	233.805	3.098	899	0.840
			Load	0.895	0.599	1.338	0.203	118	0.587
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	1.319	0.710	2.452	0.316	7582	0.381
			Age (Years)	0.874	0.615	1.243	0.179	796	0.454
	Micro-cycle ACWR 1:3-period	Next micro-cycle	Intercept	0.141	0.002	12.674	2.293	1448	0.393
			Load	1.125	0.716	1.769	0.230	771	0.609
			Sex Female (Ref)	-	-	-	-	-	-
			Sex Male	0.908	0.567	1.453	0.240	1552	0.686
			Age (Years)	1.002	0.773	1.300	0.133	1404	0.986

Abbreviations: ACWR, Acute: Chronic Workload Ratio; CI, 95% Confidence Intervals; df, Degrees of Freedom; Football Elite, Norwegian Premier League; OR, Odds Ratio; SE, Standard Error; sRPE, daily session Rating of Perceived Exertion.

[1]Daily ACWR 7:21-period was the 7-day acute sRPE divided by previous 21-day chronic sRPE per day; Micro-cycle ACWR 1:3-period was the 1-micro-cycle acute sRPE divided by previous 3-micro-cycle chronic sRPE per micro-cycle. A micro-cycle was defined as all recovery days after the previous match as well as the training days before the next match.

[2]Same day was injury same day as the measured load value; Next 4 days was one or more injuries during the four days after the measured load value; Next micro-cycle was one or more injuries during the micro-cycle after the micro-cycle of the measured load values.

Figures

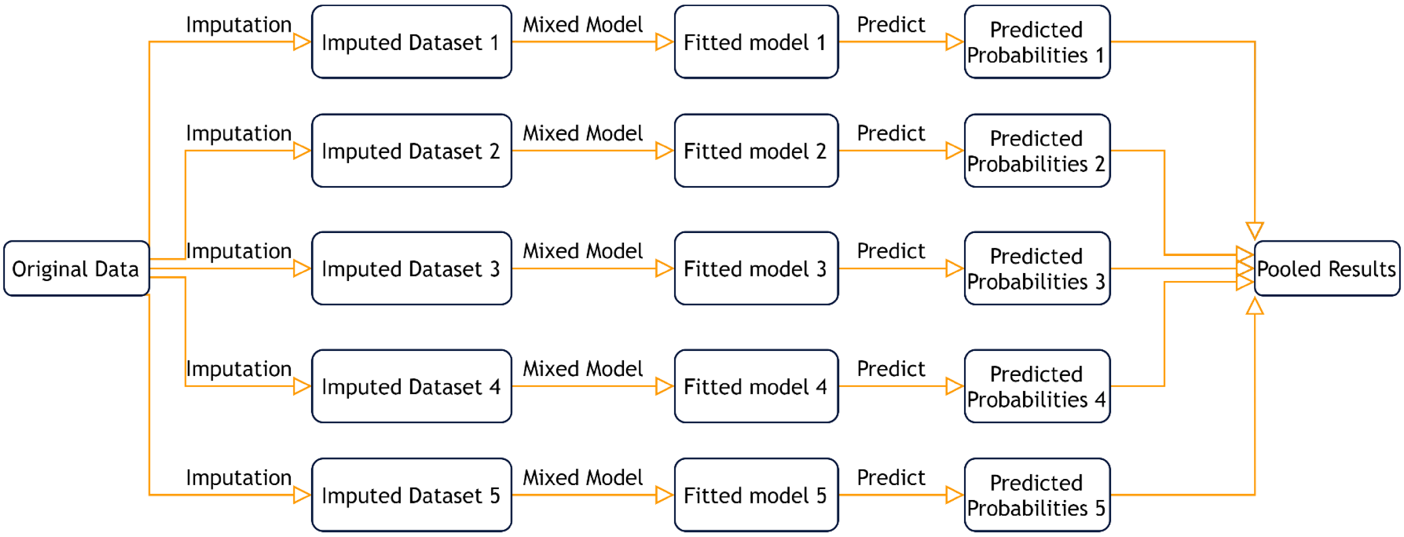

Figure S1. Illustration of the modelling process in the framework of multiple imputation. Following the recommendations in "Flexible Imputation of Missing Data, Second Edition" by Stef van Buuren,[1] which is also available online.[2] Missing load and age values were predicted and imputed using predictive mean matching.[3] All non-derived variables were used to predict imputed values, including age, sex, player position, training activity type among others. The response variable, injury, was also used to predict imputed values,[4] but was not itself imputed before analysis (guides in Van Buuren [1] 6.3.2, 6.4.1).[5] The number of imputed datasets, five, is recommended in most cases (Van Buuren section 2.8).[1] A mixed logistic regression model was run on each dataset, returning five fitted models. Each model was used to make predictions, and the mean of the predicted probabilities was used in final visualization, then the model parameters were pooled using Rubin's rules.[6]

Bache-Mathiesen LK, *et al. BMJ Open Sp Ex Med* 2021; 7:e001119. doi: 10.1136/bmjsem-2021-001119

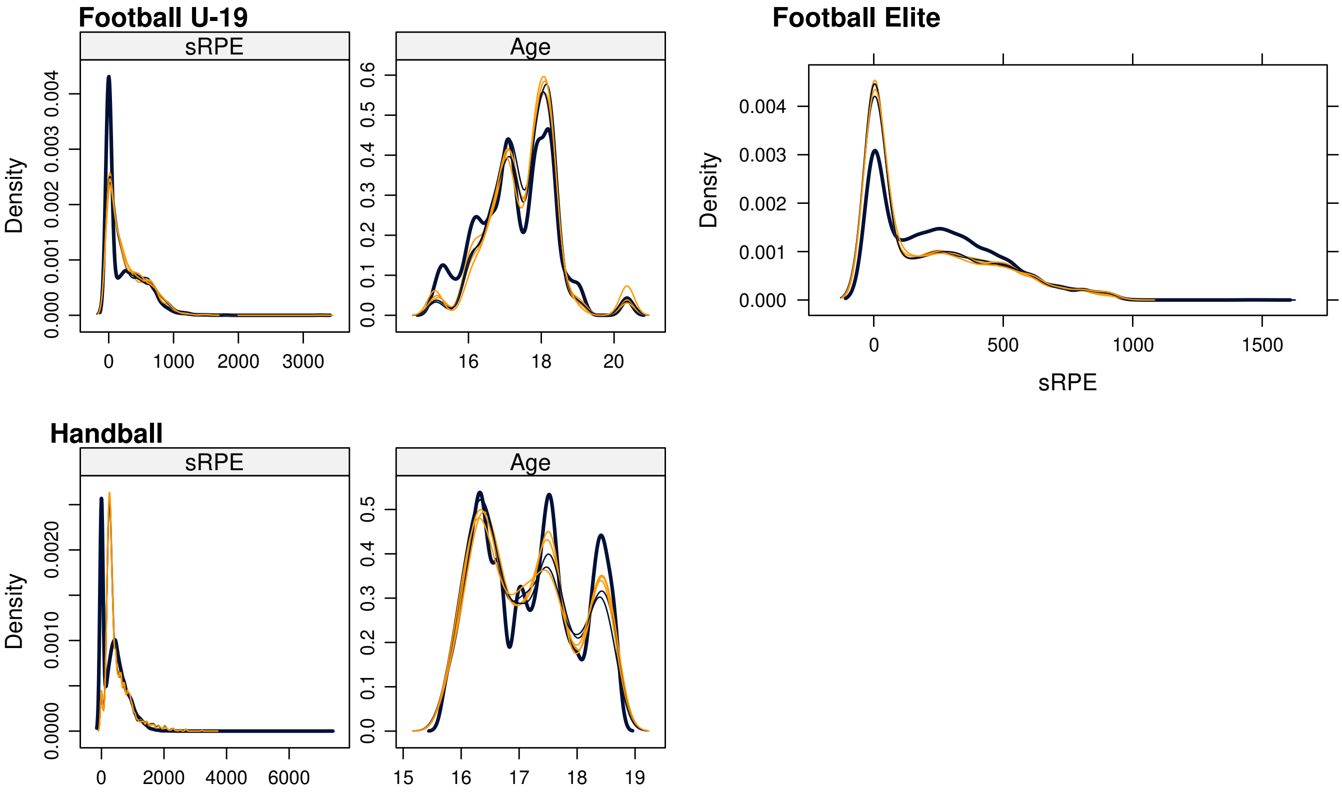

Figure S2. Distribution of real data values (blue) compared to imputed values from five imputed datasets (yellow) for the session Rating of Perceived Exertion (sRPE) measured in arbitrary units, and Age (years) in the Norwegian elite U-19 football dataset (Football U-19), the Norwegian Premier League dataset (Football Elite), and the Norwegian elite youth handball dataset. The Norwegian Premier League dataset had no missing age values.

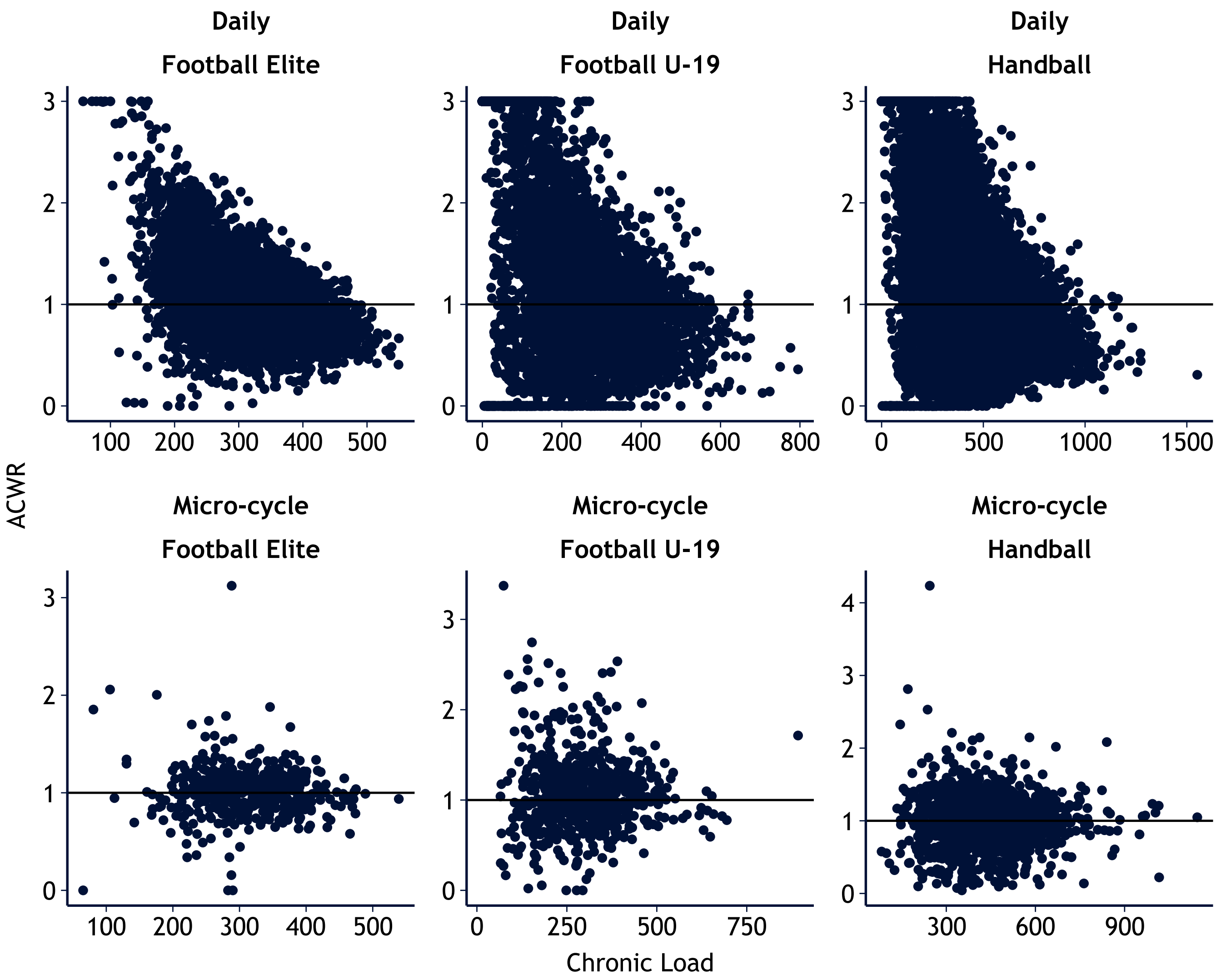

Figure S3. Scatterplot of Acute:Chronic Workload Ratio (ACWR) value vs. corresponding chronic load value (the denominator) in the Norwegian Premier League football dataset (Football Elite), the Norwegian elite U-19 football dataset (Football U-19), and Norwegian elite youth handball dataset (Handball). When computing a ratio, one assumes that there is no relationship between the ratio and the denominator after controlling for the denominator; a ratio is only effective when the relationship between the numerator and the denominator is a straight line that intersects the origin.[7] For micro-cycle ACWR, the assumption is upheld, while for daily ACWR, the assumption is violated.

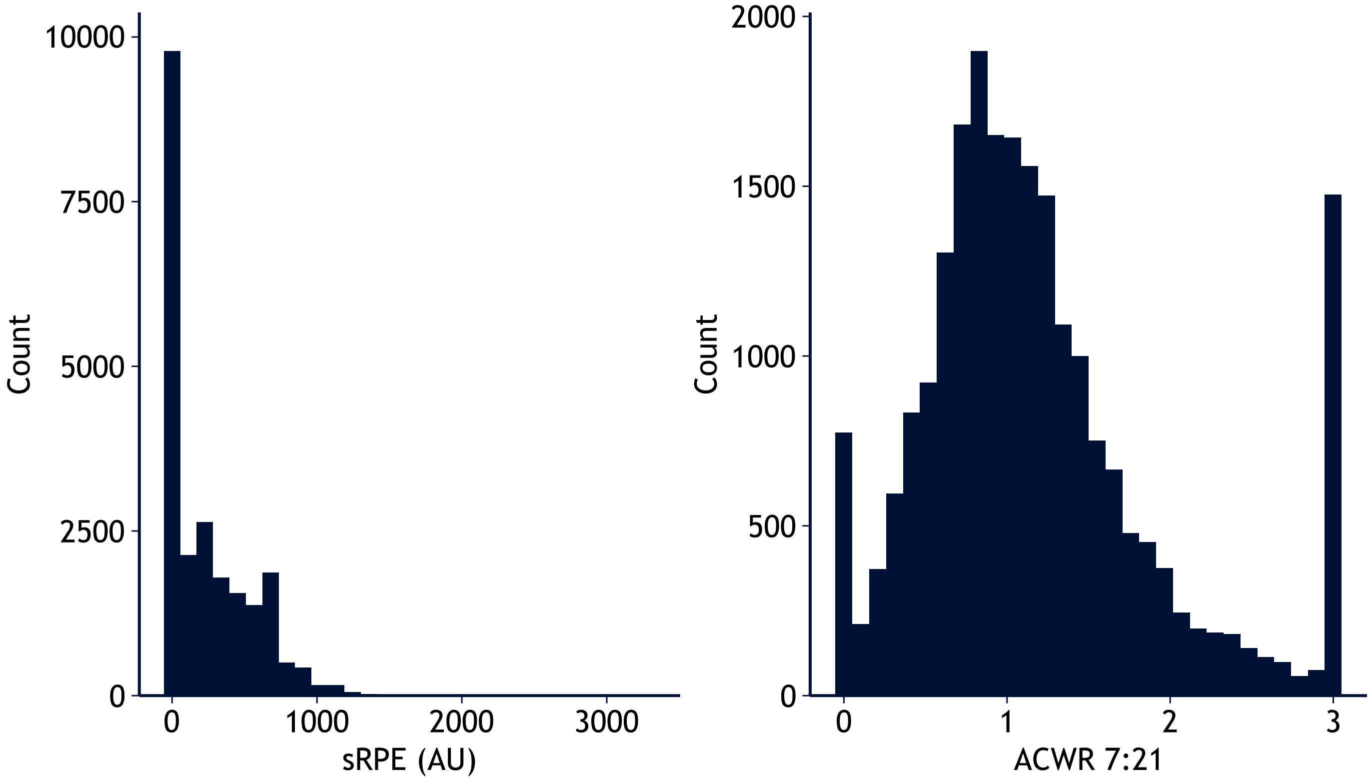

Figure S4. Distribution of the session Rating of Perceived Exertion (sRPE) reported in arbitrary units (AU), and distribution of the 7-day Acute Workload divided by 21-Chronic Workload (ACWR 7:21), from the Norwegian elite U-19 football data used as basis for simulations.

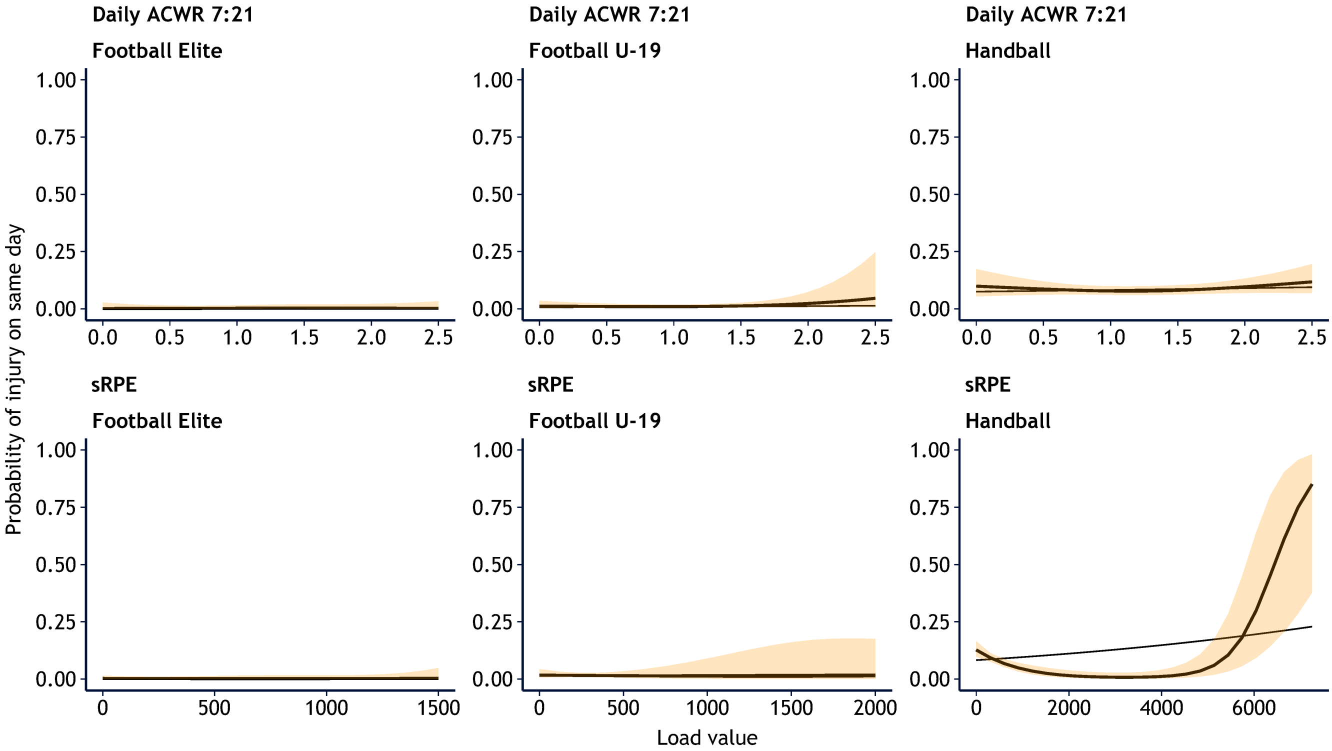

Figure S5. Probability of injury on the same day for each level of session Rating of Perceived Exertion (sRPE) and level of daily Acute:Chronic Workload Ratio (ACWR), in Norwegian Premier League (Football Elite), Norwegian elite U-19 football (Football U-19), and Norwegian elite youth handball (Handball). Probabilities are predicted by mixed-effects logistic regression models with restricted cubic splines. The yellow area represents 95% confidence intervals around predicted values. The straight line shows the same predictions from an equivalent model without splines (i.e. assuming linearity).

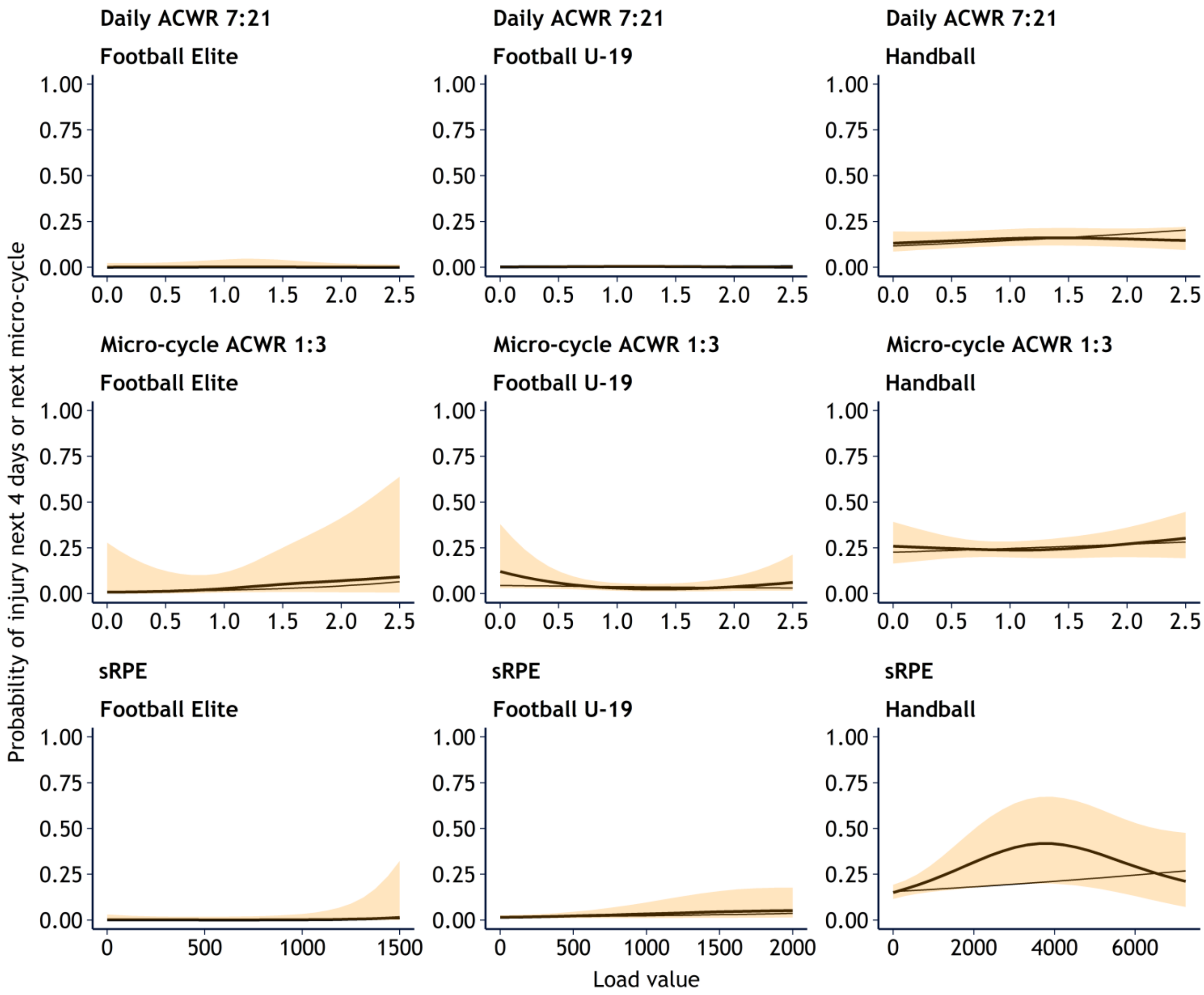

Figure S6. Probability of injury in the future for each level of daily Acute:Chronic Workload Ratio (ACWR), level of Micro-cycle ACWR, and level of session Rating of Perceived Exertion (sRPE), in Norwegian Premier League (Football Elite), Norwegian elite U-19 football (Football U-19), and Norwegian elite youth handball (Handball). Future injury was defined as any injury occurring during the next 4 days for all models except micro-cycle models, where future injury was defined as any injury occurring during the next micro-cycle. Probabilities are predicted by mixed-effects logistic regression models with restricted cubic splines. The yellow area represents 95% confidence intervals around predicted values. The straight line shows the same predictions from an equivalent model without splines (i.e. assuming linearity).

Bache-Mathiesen LK, *et al. BMJ Open Sp Ex Med* 2021; 7:e001119. doi: 10.1136/bmjsem-2021-001119

REFERENCES

1. Van Buuren S. Flexible imputation of missing data: Chapman and Hall/CRC 2018.
2. Van Buuren S. Flexible Imputation of Missing Data: CRC Press; 2018 [Second Edition:[Available from: https://stefvanbuuren.name/fimd/ accessed 2021-02-12.
3. Barzi F, Woodward M. Imputations of missing values in practice: results from imputations of serum cholesterol in 28 cohort studies. *American journal of epidemiology* 2004;160(1):34-45.
4. Moons KG, Donders RA, Stijnen T, et al. Using the outcome for imputation of missing predictor values was preferred. *Journal of clinical epidemiology* 2006;59(10):1092-101.
5. Peters SA, Bots ML, den Ruijter HM, et al. Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *Journal of clinical epidemiology* 2012;65(6):686-95.
6. Miles A. Obtaining predictions from models fit to multiply imputed data. *Sociological Methods & Research* 2016;45(1):175-85.
7. Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. *Advances in physiology education* 2013;37(3):213-19.

SUPPLEMENTARY MATERIALS AND METHODS

Participants

To find out whether the relationship between training load and injury risk may be non-linear, and whether the shape may vary between different populations, access was gained to data from different sports: football (soccer) and handball, and different populations within the same sport: Norwegian elite U-19 football data and a Norwegian Premier League football team.

The Norwegian elite U-19 data was used in Dalen-Lorentsen, et al. [1]. It was a cohort of six Norwegian elite U-19 football teams (3 female and 3 male) with 81 players (55% male, mean age: 17 years, standard deviation (SD): 1 year) followed from July to October 2017 for 104 days.

The second football cohort was a professional male football team from the Norwegian Premier League surveyed from January to December 2019 for 323 days (n = 36, mean age: 26 years (SD: 4)).[2]

The handball data was a cohort of 205 elite youth handball players from five different sport high schools in Norway (36% male, mean age: 17 years (SD: 1)) followed through a season from September 2018 to April 2019 for 237 days.[3]

Training load definition

In all three cohorts, players reported the number of training sessions and matches daily. They also reported the duration of each activity and their Rating of Perceived Exertion (RPE)[4] on the modified Borg CR10 scale.[5] To derive the session RPE (sRPE),[5] we multiplied the RPE by the activity duration in minutes. To summarize daily loads, sRPE was calculated for each session and subsequently summed.

Missing sRPE values are reported in Table S1 (Supplementary I) and were 24% for elite U-19 football, 41% for Premier League football, and 64% for elite youth handball. The values were imputed using multiple imputation, a method that also performs well in cases of high amounts of missing (80%) if the data are Missing at Random,[6] which is most common in clinical research.[7] For more detailed information on the imputation process, see Supplementary I Figure S1. The observed distribution was maintained in the imputed values; therefore the imputation was deemed valid (Figure S2).

All load measures were based on players' daily ratings of perceived exertion (sRPE). We calculated an Acute-Chronic Workload Ratio (ACWR) in two different ways:

Daily ACWR 7:21

The mean sRPE across 7 days divided by the exponentially-weighted-moving average (EWMA) of the previous 21 days (Figure 1). EWMA accounts for the assumption that load values closer in time to the event are more associated with the event than measures further back in time.[8] The calculation was uncoupled, meaning that the 7 days of acute load for the numerator were not included in the 21 days of the denominator.[9]

The calculation was performed on a sliding window moving one day at a time from and including the 28th day.[10] The last day in the acute load is considered Day 0 (Figure 1).

One limitation with the ACWR is that it bloats cases where the athlete has had little to no chronic load and returns to regular exercise. In previous studies, these cases have traditionally been deleted.[11] Here, these cases were set to have an ACWR of 3, a very high ACWR value, in line with recommendations in Harrell [12] for treatment of overly influential values. Likewise, if the EWMA chronic load was equal to zero and ACWR could not be calculated, the ACWR was set to 3.

Micro-cycle ACWR 1:3
The mean sRPE for each micro-cycle divided by the EWMA of the previous 3 micro-cycles, uncoupled (Figure 1). A micro-cycle was defined as all recovery days after the previous match and the training days before the next match. The next micro-cycle started on the first training day after the match, and so on. For an illustration of a micro-cycle, see Figure 1. The calculation was performed in the same manner as described for daily ACWR, on a sliding window moving one micro-cycle at a time from and including the 4th micro-cycle. The last day of the 4th micro-cycle was considered Day 0 (Figure 1).

When computing a ratio, one assumes that there is no relationship between the ratio and the denominator after controlling for the denominator; a ratio is only effective when the relationship between the numerator and the denominator is a straight line that intersects the origin.[13] Tests of this assumption are reported in Supplementary I Figure S3.

Injury definition
The same online questionnaire was used to collect daily health status and training information from all three sports cohorts. The elite U-19 football data and elite youth handball data were collected via the Briteback AB online survey platform, while the Norwegian Premier League football data were collected with Athlete Monitoring, Moncton, Canada.

The players daily reported whether they had experienced “no health problem”, “a new health problem”, or an “exacerbation of an existing health problem”. In the youth elite handball study, if players reported any new health problems, they were immediately prompted to specify whether it was an injury or illness in the questionnaire. In the football studies, if players reported any new health problems, a clinician contacted them by telephone the following day for a structured interview and classified the health problem as an injury or illness with the Union of European Football Associations guidelines.[14] Players were asked to report all physical complaints, irrespective of their consequences on sports participation or the need to seek medical attention.[15]

Bache-Mathiesen LK, *et al. BMJ Open Sp Ex Med* 2021; 7:e001119. doi: 10.1136/bmjsem-2021-001119

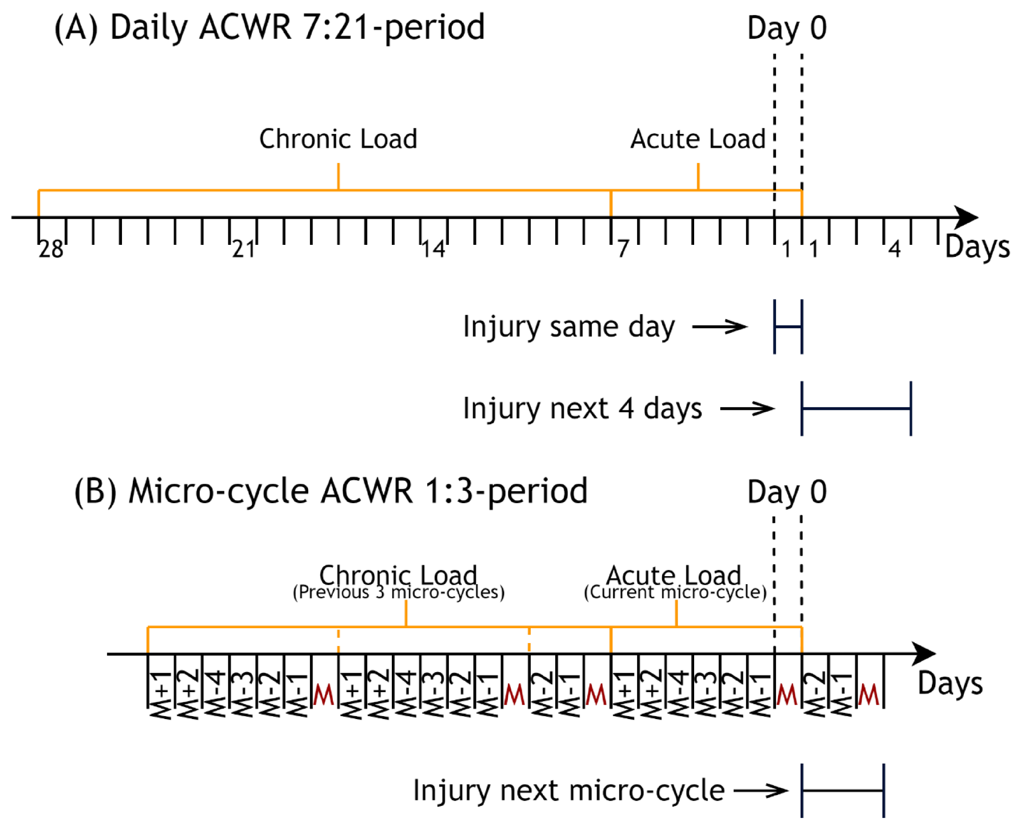

Figure 1. Illustration of time-periods for calculating (A) Daily ACWR 7:21-period and (B) Micro-cycle ACWR 1:3-period. The first day that ACWR is calculated from is denoted Day 0. The space between two tick marks represent one day (24 hours). For (B), a micro-cycle period consists of all activity before a new match (M). That is, recovery days after the previous match as well as the training days before the next match. Days denoted with negative numbers are training days before the next match (M-1; being the day before the match, M-2; two days before a match, and so on). Days with positive numbers are recovery and training days after a match (M+1; being the day after a match, M+2; two days after a match). The number of days between matches varies by the match schedule. How a team plan their training and recovery activities varies, and is dependent on the teams' philosophy. For (A), injury on the same day is defined as an injury on Day 0, and future injury is defined as an injury occurring during the next 4 days excluding Day 0. For (B) future injury was defined as an injury occurring during the next micro-cycle excluding Day 0.

Ethical Considerations

Data collection for all three studies were approved by the Ethical Review Board of the Norwegian School of Sport Sciences. They were also approved by the Norwegian Centre for Research Data: Norwegian elite U-19 football (5487); Norwegian Premier League football (722773); Norwegian elite youth handball (407930). The Norwegian elite U-19 football study was also approved by the South-Eastern Norway Regional Committee for Medical and Health Research Ethics (2017/1015). Ethical principles were followed in accordance with the Declaration of Helsinki.[16] All participants provided informed consent. All participants were above the age of 15 and parental consent was not required. Participants were assured their responses would only be available to the research team, participation was voluntary, and consent could be withdrawn at any time.

Legality of using the data in this study was dependent on the "purposes of the processing for which the personal data were intended" as written in the consent forms.[17] The consent forms for the football studies were general enough that use in this study were within the posted aims. For the elite youth handball data, the Norwegian Centre for Research Data deemed the aims described in the consent forms invalid for use in this study, and the data had to be anonymised. Anonymisation was performed under guidelines outlined by The Norwegian Data Protection Authority.[18]

Statistical analyses

To estimate the relationship between training load and injury risk, mixed-effects logistic regression was used. Logistic regression is the most frequent regression analysis in the field of training load and injury.[19] Mixed models have been recommended to account for within-player dependencies[20] and are robust to missing data in the outcome variable.[21]

All injuries were considered an event in the response variable. Illnesses and explicit replies of "no health problem" were considered non-events. Non-responses were recorded as missing. Independence between subsequent injuries within the same player was assumed.

We considered two outcomes: (1) occurrence of an injury on the same day as the observed training load (Day 0); (2) occurrence of injury in the future, where the current observation day (Day 0) was not included. For unmodified training load values and daily ACWR 7:21-period, future injury was defined as an injury occurring during the next four days excluding Day 0. For micro-cycle ACWR 1:3-period, the future injury was any injury occurring during the next micro-cycle excluding Day 0. See Figure 1 for an illustration of injury time periods and Table S2 (Supplementary I) for a list of the different models.

For models where the injury definition was set to the future, any number of injuries sustained during the time window were aggregated to 1 event. Furthermore, injuries sustained before the first calculated ACWR value had to be discarded. Consequentially, the number of injuries included in the different models varied (Table S2).

We adjusted for player age in all analyses. In addition, we adjusted for sex in the U-19 elite football and the elite youth handball models. Akaike's Information Criterion (AIC) was used to determine the model fit between including a random intercept only vs. including a random intercept & random slope for training load per player, where the best fit was chosen for the final model. Overly influential observations – extreme outliers which affect analyses – were checked using dfbeta.[12]

In all models, the relationship between sRPE and injury risk was modelled with Restricted Cubic Splines (RCS).[22] The number of knots was decided using AIC. The models were repeated without splines to simulate the relationship we would have discovered if we had assumed linearity. When using RCS, the estimated regression coefficients do not have a clinically meaningful interpretation, and only their p-values are numerically interpretable.[12] The main result is therefore a visualization of the model predictions (with 95% cluster-robust confidence intervals) to determine the shape of the relationship between training load and injury risk. To limit the number of figures to the most relevant, only predictions from models that showed a tendency towards a relationship or stronger are included in the article itself, but figures for all relationships are shown in Supplementary I Figure S5–S6. For each model, predicted values were estimated on each imputed dataset, and then pooled before visualization (Figure S1).[23]

Our analyses served to illustrate whether there is any evidence for non-linearity in training load and injury research and should not be interpreted as causal inference.

Simulation

Step 1 Preparing data

In addition to analysing real data, we performed (stochastic) simulations to compare different methods for ascertaining non-linear and linear relationships between training load and injury risk. The methodology here is focused on a causal research setting; however, the methods may also be applied in predictive research.[25] The simulations were based on the elite U-19 football dataset since it had the least missing data (24%). An imputed dataset was chosen from the 5 datasets previously imputed with multiple imputation.

Two datasets were created. The first kept the original 8 495 sRPE and 6 308 ACWR values. In the second, sRPE and ACWR were sampled with replacement to generate a scenario of 3 football teams (75 players) followed meticulously for a season (300 days), altogether 22 500 training load values. The distribution of the real data was retained during sampling; highly skewed for sRPE and Gaussian for ACWR (Figure S4).

Step 2 Generating predetermined relationships

Artificial injuries were simulated and added to each dataset under different relationship scenarios with training load. The risk models were based on the logistic function:

$$logistic(x) = \frac{1}{1 + \exp(-x)}$$

U shape

A symmetrical U parabola coinciding with the theory in Gamble 2013.[24] Using the logistic function above, the U shape function was:

$$Prob\{Y = 1|sRPE = \ logistic(-1 + \ 0.0000002 * (sRPE - 1500)^2)$$

Where Y is an indicator variable for injury.

J shape

The J shape was chosen to reproduce findings in Carey, et al. [25] with the risk function:

$$Prob\{Y = 1|ACWR\} = logistic(\begin{cases} -3.4 + 2 \cdot (1 - ACWR)^2, & ACWR < 1 \\ -3.4 + (1 - ACWR)^2, & 1 \leq ACWR < 1.7 \\ 1.5 \cdot ACWR - 5.4, & ACWR \geq 1.7 \end{cases})$$

Linear shape
A linear shape to determine whether a method optimal for non-linear modeling can also model a linear shape. The function was then:

$$Prob\{Y = 1|sRPE\} = logistic(-0.5 + 0.001 * sRPE)$$

For the U shape and linear shape, the simulated probability of an injury was based on the sRPE, while for the J shape, it was based on the ACWR.

We assumed a longitudinal design for the simulation, and an autoregressive correlation structure was implemented to ensure that values closer in time were more highly correlated than values further apart.[8] Any reference to the "true" probability refers to the simulated probability we have created for a given scenario, and which we aim to model.

While shown to be valid and reliable, the sRPE may still have some measurement error.[26] Before analyses, noise was added to load values to simulate this. The amount was set to the default jitter value, which was:

$$\frac{\max(load) - \min(load)}{50}$$

Step 3 Running models on all combinations of datasets and relationship shapes

In the same manner as in the analysis of the real data, a logistic regression model with random effects (mixed model) was used to determine the relationship between training load and predefined injury risk. Different methods of modifying training load were compared.

Linear Model
A standard logistic regression served as an example of a method which assumes linearity and illustrated the degree of error should the linearity assumption be ignored in cases where the relationship is non-linear. The purpose was to determine whether more complicated or time-consuming methods were worth the effort.

A logistic regression model describes the relationship between the probability of an event in the response variable Y (injury), given the status of the explanatory variables $X = \{x_1, x_2, \dots, x_n\}$ as the additive contribution of the intercept β_0 and linear slopes $\beta_1, \beta_2, \dots, \beta_n$ of said variables.[27] In a logistic regression with a single explanatory variable (covariate) x_1, representing the load variable, the formula is as follows:

$$Prob\{Y = 1|X\} = \frac{\exp(\beta_0 + \beta_1 x_1 + \gamma)}{1 + \exp(\beta_0 + \beta_1 x_1 + \gamma)} = logistic(\beta_0 + \beta_1 x_1 + \gamma)$$

Where γ is the random effect term.

Categorization
Although categorizing the load variable into groups before performing the intended analysis has previously been shown to be a poor method for modelling non-linear relationships,[25] we chose nevertheless to include it in our comparison of methods. For one, the method has

been recommended since.[28 29] For another, as the authors requested, we attempted to reproduce the results in another sport population under different conditions. Here, the sRPE data are highly skewed. We also increased the number of permutations for more accurate results.

To show how results may differ depending on how variables are categorized, we categorized the training load variable in two ways, before including them in two separate logistic regression models. The first was a categorization by quartiles to exemplify a data-driven approach, a chosen method in numerous studies in the past.[30-32] The second was subjectively chosen cut-offs based on the range of the data. For sRPE, four categories were made: <= 499, 500–1 499, 1 500–2 499 and >= 2 500. For ACWR, three categories were made: < 1, 1–1.74 and >= 1.75, which are the same used in Carey, et al. [25].

Quadratic model

Quadratic regression has seen some use in recent years.[33] In some studies, a quadratic term was added to the regression model to test for linearity.[34 35] Where as in others, the researchers hypothesized a parabolic shape and used quadratic regression to model the training load and injury relationship accordingly.[10 36] In a quadratic model, a polynomial to the second power is added to the standard regression model. For the logistic regression, it is denoted thus:

$$Prob\{Y = 1|X\} = \ logistic(\beta_0 + \ \beta_1 x_1 + \ \beta_2 {x_1}^2 + \gamma)$$

The model will then fit a parabolic shape between the probability of an event in Y (injury) and the explanatory variable x_1 (training load). A polynomial term can be added regardless of whether it is a linear, logistic or Poisson regression model. Although easy-to-use and intuitive, the main disadvantage of quadratic regression is that it can only model a parabola; for instance, it cannot uncover a sigmoidal shape.

Fractional polynomials

Quadratic regression is a sub-method of the more flexible Fractional Polynomials (FP), which has been used in one single training load and injury risk study.[37] Fractional polynomials, simply put, uses polynomial transformations to estimate the association between the covariate and the outcome.[38] FPs can model multiple shapes, not just the parabola. Fractional polynomials add either a single polynomial term to the pth power to the regression model (known as an FP1 model), or two polynomial terms to the pth power to the model (FP2 model).[38] The FP2 model has been shown to be the optimal choice in most cases and was chosen for all models in this study.[39] The logistic regression model with FP2 is as follows:

$$Prob\{Y = 1|X\} = \ logistic(\beta_0 + \ \beta_1 x_1 + \ \beta_2 {x_1}^{p1} + \beta_3 {x_1}^{p2} + \gamma)$$

Where $p1$ and $p2$ are exponents selected from {-2, -1, -0.5, 0, 0.5, 1, 2, 3}. A form of backward elimination was used to determine the polynomial powers with the best fit, see Ambler and Benner [40] for more details. A step-by-step guide to perform FP in R can be accessed on the primary author's GitHub.[41]

Restricted cubic splines
Another possible approach to model non-linear relationships is to use Restricted Cubic Splines (RCS). This approach as well as FP, performed better than categorization in the study by Carey, et al. [25], who found no distinct differences between RCS and FP. In cubic splines, the X-axis is divided into intervals by a number of endpoints (knots). At these knots, different cubic polynomials are joined and forced to have a consistent function, slope and acceleration (second derivative) until the next knot. At the knot, the rate change of acceleration (third derivative) may change. For three knots a, b and c, our logistic regression formula becomes:

$$Prob\{Y = 1|X\} =$$
$$logistic[\beta_0 + \beta_1 x_1 + \beta_2 {x_1}^2 + \beta_3 {x_1}^3 + \beta_4 (x_1 - a)^3 + \beta_5 (x_1 - b)^3 + \beta_6 (x_1 - c)^3 + \gamma]$$

In restricted cubic splines, the function is restricted to behave linearly in the tails.[22]

RCS has the advantage of flexibility, but the effect sizes are difficult to interpret, and the number and location of knots must be chosen, either by a data-driven or approach or as a choice of the user. As 3–5 knots are appropriate for most datasets,[12] 3 knots were used in all simulation models. We compared two different ways of choosing knot location. In the first, the knot locations were chosen by the default approach in the statistical software (data-driven), and in the other, knot locations were cut-off subjectively at sRPE = 500, 1 500 and 2 500, and likewise at ACWR = 1, 1.75 and 2, to cover the range of the load metrics.

A step-by-step guide to perform RCS in R can be accessed on the primary author's GitHub.[42]

Step 4 Calculating performance metrics

The Root-Mean-Squared Error (RMSE) was calculated to numerically evaluate the accuracy of the methods. RMSE is a combined measure of accuracy and precision, where the lower the RMSE, the better the method. RMSE was calculated as the square root of the mean difference between the true risk and predicted risk for each observation. The scale of the RMSE depends on the analysis in question, and it is therefore only interpretable by comparing values in the same analysis – the values cannot be interpreted in isolation.[43]

To supplement RMSE, the proportion of prediction intervals that included the true coefficient was calculated (coverage). Brier score for model fit and C-statistics (also known as the concordance, or as the area under the receiving operating characteristic curve) was calculated for predictive ability, since they are commonly used in training load and injury risk studies.[44-47]

Final analyses

In summary, the four steps of the simulation were:

1 Sample training load values from the elite U-19 football data
2 Simulate injuries with three different shapes for the relationship between injury risk and training load

3 Fit seven different models with injury as the outcome and training load as the explanatory variable
4 Calculate performance measures

Using formulas listen in Morris, et al. [43], accepting a Monte Carlo Standard Error of no more than 0.5, the number of permutations needed for an accurate determination of coverage was:

$$n_{Coverage} = \frac{E(Coverage)(1 - E(Coverage)}{(Monte\ Carlo\ SE_{req})^2} = \frac{95 * 5}{0.5^2} = 1\ 900$$

Steps 1–4 were therefore repeated 1 900 times for all relationship scenarios.

For the U-shaped relationship, predicted values were visualized alongside the predefined shape to determine each method's ability to capture the true relationship. Only one permutation was used for the visualization to avoid cluttering of lines.

The mean RMSE, coverage, C-statistics and Brier score were calculated for each combination of model-method and dataset sizes for the U-, J- and linear-shaped relationships. As mean RMSE was the most relevant metric for determining model accuracy, it was visually compared for the non-linear shapes.

All statistical analyses and simulations were performed using R version 4.0.2[48] with RStudio version 1.3.1056. Packages were used for specific purposes: multiple imputation with MICE,[49] mixed models with lme4,[50] predictions with ggeffects,[51] confidence intervals with clubSandwich,[52] predictions with prediction intervals using merTools,[53] and splines with the rms package.[54] The simulations were run on a computer with an Intel(R) Core(TM) i7-6700K 4.00GHz CPU, and with 16 GB RAM. A GitHub repository is available with all R code and the data used in the simulations.[55]

REFERENCES

1. Dalen-Lorentsen T, Andersen TE, Bjørneboe J, et al. A cherry tree ripe for picking: The relationship between the acute: chronic workload ratio and health problems. *Journal of Orthopaedic & Sports Physical Therapy* 2021(0):1-26.
2. Theron GF. The use of Data Mining for Predicting Injuries in Professional Football Players. University of Oslo, 2020.
3. Bjørndal CT, Bache-Mathiesen LK, Gjesdal S, et al. An Examination of Training Load, Match Activities, and Health Problems in Norwegian Youth Elite Handball Players Over One Competitive Season. *Frontiers in Sports and Active Living* 2021;3(36) doi: 10.3389/fspor.2021.635103
4. Borg G, Hassmén P, Lagerström M. Perceived exertion related to heart rate and blood lactate during arm and leg exercise. *European journal of applied physiology and occupational physiology* 1987;56(6):679-85.
5. Foster C, Florhaug JA, Franklin J, et al. A New Approach to Monitoring Exercise Training. *Journal of Strength and Conditioning Research* 2001;15:10Y 115.
6. Madley-Dowd P, Hughes R, Tilling K, et al. The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of clinical epidemiology* 2019;110:63-73.

7. Janssen KJ, Donders ART, Harrell Jr FE, et al. Missing covariate data in medical research: to impute is better than to ignore. *Journal of clinical epidemiology* 2010;63(7):721-27.
8. Williams S, West S, Cross MJ, et al. Better way to determine the acute: chronic workload ratio? *British journal of sports medicine* 2017;51(3):209-10.
9. Wang C, Vargas JT, Stokes T, et al. Analyzing Activity and Injury: Lessons Learned from the Acute: Chronic Workload Ratio. *Sports Medicine* 2020:1-12.
10. Carey DL, Blanch P, Ong K-L, et al. Training loads and injury risk in Australian football—differing acute: chronic workload ratios influence match injury risk. *British journal of sports medicine* 2017;51(16):1215-20.
11. Hulin BT, Gabbett TJ, Lawson DW, et al. The acute: chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *Br J Sports Med* 2016;50(4):231-36.
12. Harrell FE. Regression modeling strategies. *BIOS* 2017;330:2018.
13. Curran-Everett D. Explorations in statistics: the analysis of ratios and normalized data. *Advances in physiology education* 2013;37(3):213-19.
14. Hägglund M, Waldén M, Bahr R, et al. Methods for epidemiological study of injuries to professional football players: developing the UEFA model. *British journal of sports medicine* 2005;39(6):340-46.
15. Fuller CW, Ekstrand J, Junge A, et al. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Scandinavian journal of medicine & science in sports* 2006;16(2):83-92.
16. Malik AY, Foster C. The revised Declaration of Helsinki: cosmetic or real change? *Journal of the Royal Society of Medicine* 2016;109(5):184-89.
17. EU P. General Data Protection Agency Article 13: intersoft consulting; 2018 [Available from: https://gdpr-info.eu/art-13-gdpr/ accessed 2021-02-10.
18. Datatilsynet. The anonymisation of personal data: Norwegian Data Protection Agency; 2017 [Available from: https://www.datatilsynet.no/en/regulations-and-tools/reports-on-specific-subjects/anonymisation/ accessed 2021-04-12.
19. Windt J, Ardern CL, Gabbett TJ, et al. Getting the most out of intensive longitudinal data: a methodological review of workload–injury studies. *BMJ open* 2018;8(10):e022626.
20. Nielsen RO, Shrier I, Casals M, et al. Statement on Methods in Sport Injury Research From the First METHODS MATTER Meeting, Copenhagen, 2019. *Journal of Orthopaedic & Sports Physical Therapy* 2020;50(5):226-33.
21. Peters SA, Bots ML, den Ruijter HM, et al. Multiple imputation of missing repeated outcome measurements did not add to linear mixed-effects models. *Journal of clinical epidemiology* 2012;65(6):686-95.
22. Durrleman S, Simon R. Flexible regression models with cubic splines. *Statistics in medicine* 1989;8(5):551-61.
23. Miles A. Obtaining predictions from models fit to multiply imputed data. *Sociological Methods & Research* 2016;45(1):175-85.
24. Gamble P. Reducing injury in elite sport-is simply restricting workloads really the answer? *New Zealand Journal of Sports Medicine* 2013:36.
25. Carey DL, Crossley KM, Whiteley R, et al. Modeling Training Loads and Injuries: The Dangers of Discretization. *Medicine and science in sports and exercise* 2018;50(11):2267-76.
26. Haddad M, Stylianides G, Djaoui L, et al. Session-RPE method for training load monitoring: validity, ecological usefulness, and influencing factors. *Frontiers in neuroscience* 2017;11:612.
27. Silk MJ, Harrison XA, Hodgson DJ. Perils and pitfalls of mixed-effects regression models in biology. *PeerJ* 2020;8:e9522.
28. Ruddy JD, Cormack SJ, Whiteley R, et al. Modeling the risk of team sport injuries: a narrative review of different statistical approaches. *Frontiers in physiology* 2019;10

29. Nielsen RO, Bertelsen ML, Ramskov D, et al. Time-to-event analysis for sports injury research part 1: time-varying exposures. *Br J Sports Med* 2019;53(1):61-68.
30. Cross MJ, Williams S, Trewartha G, et al. The influence of in-season training loads on injury risk in professional rugby union. *International journal of sports physiology and performance* 2016;11(3):350-55.
31. Malone S, Hughes B, Doran DA, et al. Can the workload–injury relationship be moderated by improved strength, speed and repeated-sprint qualities? *Journal of science and medicine in sport* 2019;22(1):29-34.
32. Stares J, Dawson B, Peeling P, et al. Identifying high risk loading conditions for in-season injury in elite Australian football players. *Journal of science and medicine in sport* 2018;21(1):46-51.
33. Gabbett TJ. The training—injury prevention paradox: should athletes be training smarter and harder? *British journal of sports medicine* 2016;50(5):273-80.
34. Sampson JA, Murray A, Williams S, et al. Injury risk-workload associations in NCAA American college football. *Journal of science and medicine in sport* 2018;21(12):1215-20.
35. Warren A, Williams S, McCaig S, et al. High acute: chronic workloads are associated with injury in England & Wales Cricket Board Development Programme fast bowlers. *Journal of science and medicine in sport* 2018;21(1):40-45.
36. Blanch P, Gabbett TJ. Has the athlete trained enough to return to play safely? The acute: chronic workload ratio permits clinicians to quantify a player's risk of subsequent injury. *British journal of sports medicine* 2016;50(8):471-75.
37. Lathlean TJ, Gastin PB, Newstead SV, et al. Absolute and Relative Load and Injury in Elite Junior Australian Football Players Over 1 Season. *International Journal of Sports Physiology and Performance* 2019;1(aop):1-9.
38. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 1994;43(3):429-53.
39. Binder H, Sauerbrei W, Royston P. Comparison between splines and fractional polynomials for multivariable model building with continuous covariates: a simulation study with continuous response. *Statistics in Medicine* 2013;32(13):2262-77.
40. mfp: Multivariable Fractional Polynomials. R package version [program], 2015.
41. Bache-Mathiesen L. Performing Fractional Polynomials in R [PDF]. GitHub; 2021 [Available from: https://github.com/lenakba/load-injury-non-linearity-study/blob/main/guide_fractional_polynomials.pdf accessed 2021-04-12.
42. Bache-Mathiesen L. Performing Restricted Cubic Splines in R [PDF]. GitHub; 2021 [Available from: https://github.com/lenakba/load-injury-non-linearity-study/blob/main/guide_restricted_cubic_splines.pdf accessed 2021-04-12.
43. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Statistics in medicine* 2019;38(11):2074-102.
44. Johnston R, Cahalan R, Bonnett L, et al. Training load and baseline characteristics associated with new injury/pain within an endurance sporting population: a prospective study. *International journal of sports physiology and performance* 2019;14(5):590-97.
45. McCall A, Dupont G, Ekstrand J. Internal workload and non-contact injury: a one-season study of five teams from the UEFA Elite Club Injury Study. *British journal of sports medicine* 2018;52(23):1517-22.
46. Thornton HR, Delaney JA, Duthie GM, et al. Importance of various training-load measures in injury incidence of professional rugby league athletes. *International journal of sports physiology and performance* 2017;12(6):819-24.
47. Visnes H, Bahr R. Training volume and body composition as risk factors for developing jumper's knee among young elite volleyball players. *Scandinavian journal of medicine & science in sports* 2013;23(5):607-13.
48. Team RC. 2021.

49. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1-67. [program]. URL https://www.jstatsoft.org/v45/i03/. version, 2011.
50. Bates D, Maechler M, Bolker B, et al. Fitting Linear Mixed-Effects Models Using lme4. . *Journal of Statistical Software, 67(1), 1-48* 2015 doi: doi:10.18637/jss.v067.i01.
51. Lüdecke D. ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software* 2018;3(26):772.
52. clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections. [program]. 0.5.3 version, 2021.
53. merTools: Tools for Analyzing Mixed Effect Regression Models. R package version 0.5.0. [program]. https://CRAN.R-project.org/package=merTools version, 2019.
54. rms: Regression Modeling Strategies. R package version 5.1-4. [program]. https://CRAN.R-project.org/package=rms version, 2019.
55. Bache-Mathiesen L. Training load and injury non-linearity R code GitHub: GitHub; 2021 [Available from: https://github.com/lenakba/load-injury-non-linearity-study accessed 2021-04-12.