# 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

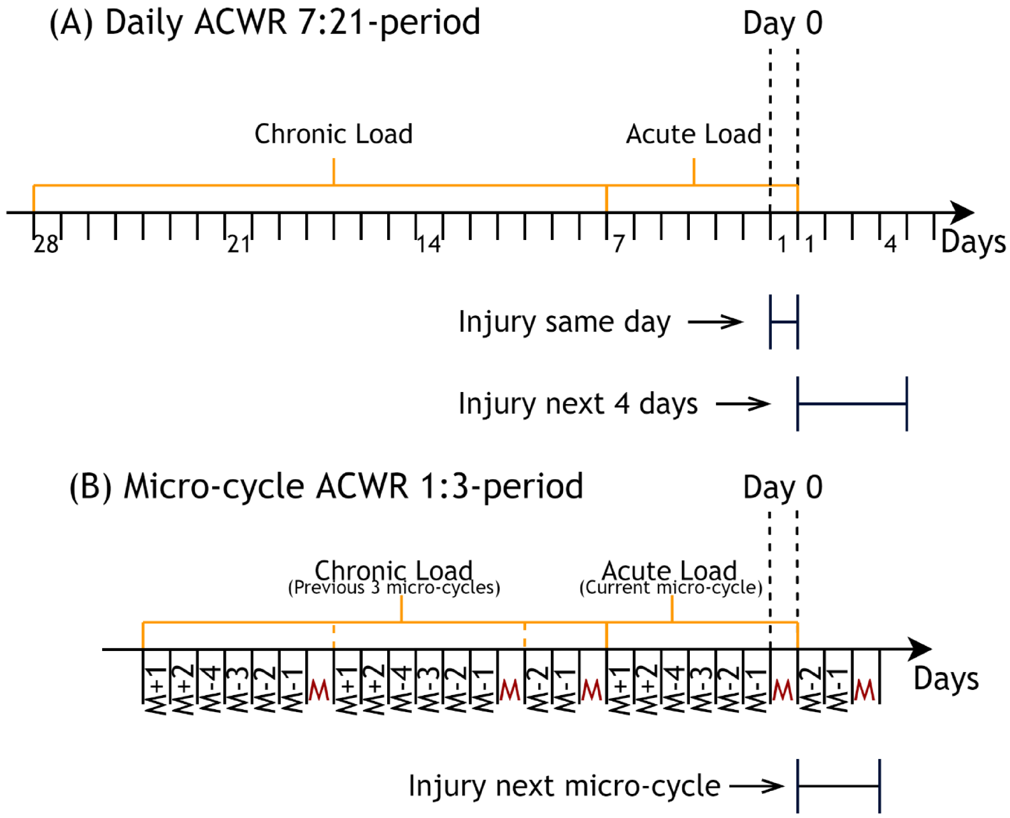


**Figure 1.** Illustration of time-periods for calculating (A) Daily ACWR 7:21-period and (B) Micro-cycle ACWR 1:3-period. The first day that ACWR is calculated from is denoted Day 0. The space between two tick marks represent one day (24 hours). For (B), a micro-cycle period consists of all activity 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.



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

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



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

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 $\beta_0$ 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 $\gamma$ 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.