BMJ Open Sport & Exercise Medicine

Assessing the cumulative effect of long-term training load on the risk of injury in team sports

Lena Kristin Bache-Mathiesen [1], Thor Einar Andersen,[2] Torstein Dalen-Lorentsen [2,3] Benjamin Clarsen,[2,4] Morten Wang Fagerland[2,5]

To cite: Bache-Mathiesen LK, Andersen TE, Dalen-Lorentsen T, *et al*. Assessing the cumulative effect of long-term training load on the risk of injury in team sports. *BMJ Open Sport & Exercise Medicine* 2022;**8**:e001342. doi:10.1136/bmjsem-2022-001342

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjsem-2022-001342).

Accepted 16 May 2022

Check for updates

© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

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

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

ABSTRACT

Objectives Determine how to assess the cumulative effect of training load on the risk of injury or health problems in team sports.

Methods First, we performed a simulation based on a Norwegian Premier League male football dataset (n players=36). Training load was sampled from daily session rating of perceived exertion (sRPE). Different scenarios of the effect of sRPE on injury risk and the effect of relative sRPE on injury risk were simulated. These scenarios assumed that the probability of injury was the result of training load exposures over the previous 4 weeks. We compared seven different methods of modelling training load in their ability to model the simulated relationship. We then used the most accurate method, the distributed lag non-linear model (DLNM), to analyse data from Norwegian youth elite handball players (no. of players=205, no. of health problems=471) to illustrate how assessing the cumulative effect of training load can be done in practice.

Results DLNM was the only method that accurately modelled the simulated relationships between training load and injury risk. In the handball example, DLNM could show the cumulative effect of training load and how much training load affected health problem risk depending on the distance in time since the training load exposure.

Conclusion DLNM can be used to assess the cumulative effect of training load on injury risk.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Training load seems to affect the risk of injury in team sports.
⇒ Time since exposure to training load may determine the strength and the direction of training load's effect on injury risk.
⇒ The ability of current methodology to assess abovementioned effects is limited.

WHAT THIS STUDY ADDS

⇒ Distributed lag non-linear models (DLNMs) were superior to all methods compared and could determine the cumulative effect of past training load.
⇒ The exponentially weighted moving average (EWMA) performed better than the rolling average and robust exponential decreasing index.
⇒ The difference between the acute:chronic workload ratio and week-to-week percentage change was negligible.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Researchers can estimate the effects of training load on the risk of injury in team sports using DLNM.
⇒ More consistent methodology in training load and injury risk studies will improve comparability and reproducibility.

INTRODUCTION

In recent years, researchers have attempted to determine the effect of training load on the risk of sports injuries and other sports-related health problems.[1] Training load is the physical exertion that the athlete has been exposed to and is a combination of the exposure itself (external load) and the physiological and psychological stressors applied to the athlete in response to the exposure (internal load).[2] Relationships between risk factors and sports injuries are often complex,[3] as the effect of risk factors may depend on the presence or absence of other risk factors,[3] the current state of the athlete,[4] and they may also act non-linearly on the risk of injury.[4]

Assessing training load poses additional challenges.[5 6] It is a multidimensional construct that can be measured in multiple ways.[7] Hypotheses suggest that not only the amount of training load, but also the relative change in training load affect injury risk.[5] Balanced training load exposure may both cause and protect against injury through building fitness and fatigue.[8] A central concern is that training load is a time-varying exposure with special properties.[5 9] The training load exposure on the current day affects injury risk directly—an athlete cannot sustain a sports injury without participating in a sporting activity.[5] Training load may, however, also be a so-called time-lagged effect.[10] The training load on the previous day may contribute to the injury risk on the

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

current day. To further add complexity, training load is likely to have a *protracted* time-lagged effect.[11] The injury risk at any given time is the result of multiple training load exposure events of different intensities sustained in the past.[12] In summary, no single training load exposure event is thought to affect injury risk in isolation, rather it is the long-term exposure to training load leading up to the event collectively that is assumed to influence injury occurrence.

To meet these assumptions, previous research has addressed some of the complexities of modelling training load statistically.[9 13] A statistical model is a generalisation that is unlikely to tailor the prognostic course of an individual accurately,[14] but it may inform researchers and clinicians about causation and patterns of injury risk. Among others, statistical solutions have been proposed to handle the time-varying effects,[9] the potential for non-linear effects,[15] the cumulative effect,[13 16] and the effect of relative training load[17] in the risk of injury. While statistical models and approaches have been recommended to handle these challenges in isolation, it is still unknown how to explore all the raised challenges in symphony. That is, accounting for time-varying effects, non-linear effects and cumulative effects simultaneously.

We aimed to determine how to model training load when assessing its cumulative effect on the risk of injury or health problems in a longitudinal team sports study.

MATERIALS AND METHODS

First, we ran a simulation study based on football data with internal training load measures to compare the performance of different statistical approaches. Then, we implemented the best performing approach on a handball dataset with training load and injury measures to demonstrate how it can be used in practice.

Football data simulation

To compare the performance of different statistical approaches, it is common to run stochastic simulations.[18] We constructed different relationships between training load and injury based on a dataset of Norwegian Premier League male football players followed for 323 days (n=36, mean age 26 years (min: 16, max: 34)).[19] We used seven methods to model the relationship between training load and injury risk. To compare the performance of the seven methods, we calculated the deviation between the relationship estimated by each method and the 'true' simulated relationship (box 1, online supplemental file 1, online supplemental figure S1). More details about all methods are available in online supplemental file 2.

Analyses and simulations were performed using R 4.1.2.[20–22] Code and data are available online.[23]

Box 1 Summary of the football data simulation

1. Sample session rating of perceived exertion values from observed training load data in football.
2. Simulate time-to-event relationships between training load and injury with seven different scenarios of time-dependent effects.
3. Use four different methods on the absolute training load and three different methods on the relative training load to model the relationship between training load and simulated injuries.
4. Calculate performance measures.

Step 1: preparing data

Internal training load was measured with the daily session rating of perceived exertion (sRPE)[24]: the duration of the activity in minutes multiplied by the player's reported perceived intensity of the activity on a scale from 0 to 10. We simulated a training load study by sampling sRPE values from the observed football dataset. The relative training load from 1 day to the next was calculated with the symmetrized percentage change (%ΔsRPE).[25] A larger study was simulated: 250 participants (10 football teams), followed for one full season (300 days).

Step 2: simulating time-to-event data

We simulated injuries under different relationship scenarios with the sampled training load. The risk of injury at any given time was predetermined with a time-to-event Cox regression model. Only one injury was simulated per individual. We use the term injury to describe the simulated events. However, the events can also be considered occurrences of pain or other health problems.

The relationship between absolute training load and injury risk was simulated to be J-shaped (online supplemental file 1 figure S2A).[15] Under this assumption, the lowest point of risk was intermediate levels of training load. The highest point of risk was set at high levels of training load.

For relative training load, we simulated a linear relationship with injury risk (online supplemental figure S2C). Higher loads on the current day compared with load on the previous day increased risk, while lower loads on the current day compared with the previous day reduced risk.[8]

In addition, we simulated the following time-dependent scenarios for both the absolute training load and the relative training load (online supplemental file S3):

- Constant. Across 4 weeks (28 days), the effect of training load has a constant effect each day.
- Decay. Across 4 weeks (28 days), the effect of training load gradually decays for each day.[13] This was hypothesised as a likely scenario if past training load has a direct effect on injury risk.
- Exponential decay. On the current day (day 0), training load has the highest risk of injury. The effect of training load drops exponentially the past 4 weeks (28 days). This was hypothesised as a likely scenario if past training load has an indirect effect on injury risk.
- Direct, then inverse. Training load values on the current week (acute) increases risk of injury, while the training load values 3 weeks before the current week (chronic) decreases risk of injury (results in supplementary).[17] This scenario represents a hypothesis that

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

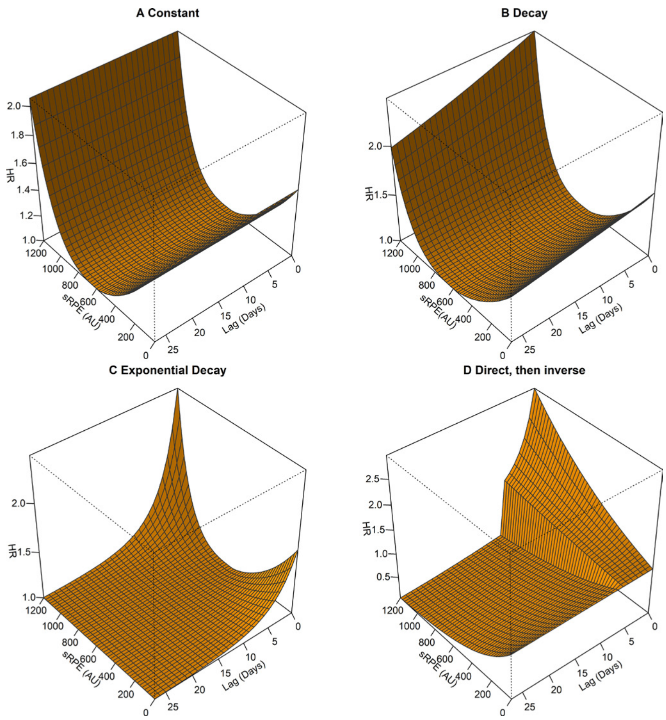

Figure 1 The four simulated relationships between absolute training load and injury risk. The relationships are a combination of the J-shaped function on the absolute training load exposure (figure 2A) and the different functions on the time since training load was sustained (online supplemental figure S3). Training load is measured with the session rating of perceived exertion (sRPE), shown on the x-axis. The time since the current day (day 0) is shown on the y-axis, where 0 is the current day and 27 is the 27th day before the current day. On the z-axis, the risk of injury is measured with the Hazard Ratio (HR), where HR >1 indicates an increased risk, and HR <1 indicates a decreased risk. The four risk shapes are: (A) constant, where the J-shaped risk of training load is constant over time; (B) decay, where the effect size of the J-shaped effect of training load is at its highest on the current day (day 0) and is reduced linearly for each lag day back in time; (C) exponential decay, where the J-shaped risk of training load is at its highest on the current day (day 0) and is reduced exponentially for each lag day back in time; (D) direct, then inverse; where training load linearly increases injury risk during the current week (day 0–6), but linearly decreases injury risk thereafter. This was the shape simulated with a linear model on the absolute training load (online supplemental figure S2B). Training load had no effect after the 27th lag day (4 weeks) in all four scenarios (not shown).

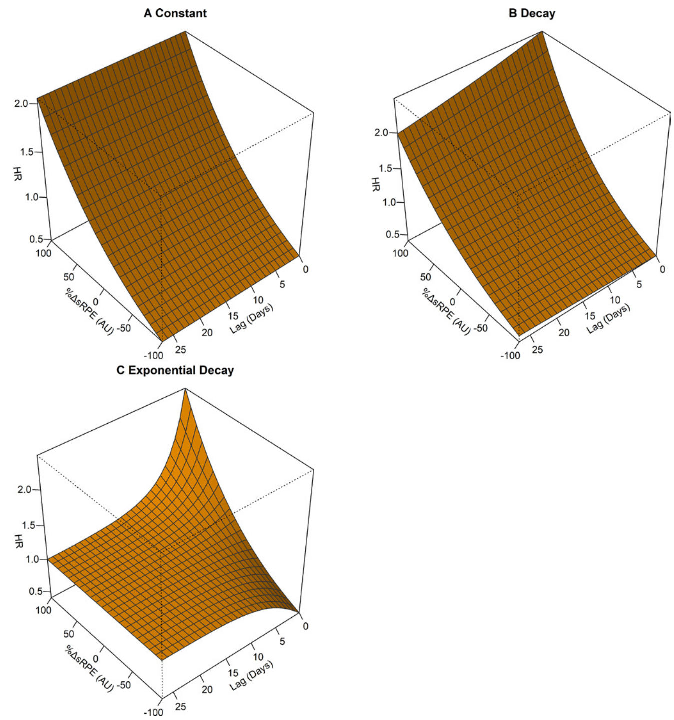

Figure 2 The three simulated relationships between relative training load and injury risk. The relationships are a combination of the linear function on the relative training load exposure (figure 2C) and the different functions on the time since training load was sustained (online supplemental figure S3). Relative training load is measured with the symmetrised percentage change (%Δ) in session rating of perceived exertion (sRPE), shown on the x-axis. The time since the current day (day 0) is shown on the y-axis, where 0 is the current day and 27 is the 27th day before the current day. On the z-axis, the risk of injury is measured with the Hazard Ratio (HR), where HR >1 indicates an increased risk, and HR <1 indicates a decreased risk. The four risk shapes are: (A) constant, where the linear risk of relative training load is constant over time; (B) decay, where the effect size of the linear effect of relative training load is at its highest on the current day (day 0) and is reduced linearly for each lag day back in time; (C) exponential decay, where the linear risk of training load is at its highest on the current day (day 0) and is reduced exponentially for each lag day back in time. Training load had no effect after the 27th lag day (4 weeks) in all three scenarios (not shown).

chronic load is a measure of fitness and absolute acute load is a measure of fatigue.[17] High loads relative to the previous time period are thought to increase risk, while low loads relative to the previous time period decrease risk: a linear relationship. Therefore, for this time-lag scenario, we simulated a linear relationship with the absolute training load, and the relative training load was not considered (online supplemental figure S2B).

In summary, seven different relationships between training load and injury risk were simulated (figures 1–2).

Step 3: modelling the time-dependent effect of training load on injury risk

Different methods of modelling training load were compared in their ability to uncover the seven predetermined relationships between training load and injury risk. We chose the most frequently used methods in training load and injury research,[26 27] methods proposed as potential alternatives[13 16] and a method developed to handle similar challenges in epidemiology.[10] Cox regression was used to estimate the relative risk of injury, where

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

internal training load, sRPE or %ΔsRPE was modified or modelled with different methods.

For absolute training load, we modelled the following methods with a quadratic term:

- Rolling average (RA).[28]
- Exponentially weighted moving average (EWMA).[13]
- Robust exponential decreasing index (REDI).[16]
- Distributed lag non-linear model (DLNM).[10 12]

For relative load, we modelled the following methods with a linear term:

- Week-to-week percentage change.[29]
- Acute:chronic workload ratio (ACWR),[17] 7:28 coupled RA.[30]
- DLNM.

Step 4: calculating performance measures to compare methods

We visualised the predicted cumulative risk versus the true cumulative risk in line graphs. The root-mean-squared-error (RMSE), a combined measure of accuracy and precision, was calculated between the predicted and true cumulative hazard. The lower the RMSE, the better the method. We also calculated RMSE on the predicted injury value versus the observed value (the model residuals).

The Akaike's Information Criterion (AIC) for model fit, coverage of 95% CI, average width of CI and the percentage of simulations where the methods had the lowest RMSE and lowest AIC were also calculated.

Implementation in a handball dataset

The model that performed best in our preliminary analyses of simulated data, the DLNM, was implemented on an actual data set from another team sport, to illustrate how it can be used in practice. To explore the potential for a time-dependent, cumulative effect of training load on health problem risk, we chose a Norwegian elite youth handball cohort (n=205, 36% male, mean age: 17 years (SD: 1), followed 237 days). Although the high amount of missing data (64% of sRPE values) renders it unsuitable for a study of causal inference, it had a sufficient number of health problems for the current methodology study (n=471 health problems).

RPE and duration were collected from the players after each training and match, from which daily sRPE was determined.[31] The handball players reported daily whether they had 'no health problem' or 'a new health problem'. Any response of 'a new health problem' was considered an event in the current study. Players were encouraged to report all physical complaints, irrespective of their consequences on sports participation or the need to seek medical attention.[32]

Missing sRPE data were imputed with multiple imputation.[33] Cox regression was run with health problem (yes/no) as the outcome and the DLNM of sRPE as the exposure.[9] We adjusted for sex and age as potential confounders and included a frailty term to account for recurrent events.[34] DLNM combines a function on the magnitude of sRPE and a function of the distance since day 0 up to lag 27 (4 weeks). The sRPE was modelled with restricted cubic splines[15] and the lag function with a linear model. The model predictions were visualised to assess the ability of DLNM to explore effects.

RESULTS

Football data simulation

Absolute training load

TheDLNM was the only method that discovered the simulated J-shaped relationship between absolute training load and cumulative risk of injury under all the main time-dependent effects (figure 3). It had, by far, the lowest mean external RMSE (online supplemental file 1 figure S4A-C), the lowest internal RMSE (table 1) and the lowest AIC (online supplemental figure S4D-F). Despite consistently having the narrowest average CI width (≈2 vs >3 (all other methods)), it also had the second-to-highest coverage of 95% CIs under the constant scenario and the highest under the decay scenario (table 1). Except for the exponential decay scenario, all methods had poor coverage overall (<=35%, table 1).

The EWMA was able to detect the exponential decay scenario (figure 3J) and had better accuracy than the rolling average and the robust exponential decreasing index for the decay scenario (figure 3E–G). It had the lowest mean external RMSE and AIC of all three scenarios and methods (table 1, online supplemental figure S4), although, under the constant scenario, the CIs reached negative values (figure 3B).

The rolling average was able to model the constant scenario (figure 3A) and had a mean internal RMSE of 0.113547, slightly lower than EWMA at 0.113548. Under this condition, it had the second best (rank 2) external RMSE in 31% of simulations and third best (rank 3) in 52% of simulations, with similar results for AIC (31% rank 2, 58% rank 3; online supplemental table S1). Here, EWMA was most frequently ranked second best for RMSE and AIC (45% and 39%, respectively (online supplemental table S1).

REDI had consistently the highest mean external RMSE and AIC (online supplemental figure S4, table 1). It was most frequently rank 4 for external RMSE under the constant and decay scenarios and for AIC under all scenarios (online supplemental table S1). Furthermore, REDI consistently had the lowest coverage of 95% CIs (table 1). Instead of discovering that high levels of absolute training load increases injury risk, REDI estimated that high absolute training load decreases injury risk under the exponential decay scenario (figure 3K).

No method was able to accurately model the direct, then inverse scenario (coverage=0%, online supplemental figure S5, online supplemental table S2).

Relative training load

The Distributed Lag Non-Linear Model (DLNM) was also capable of discovering the cumulative hazard of injury for relative training load (figure 4C, F, I). It had the lowest mean internal RMSE and AIC for the Constant and Decay

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

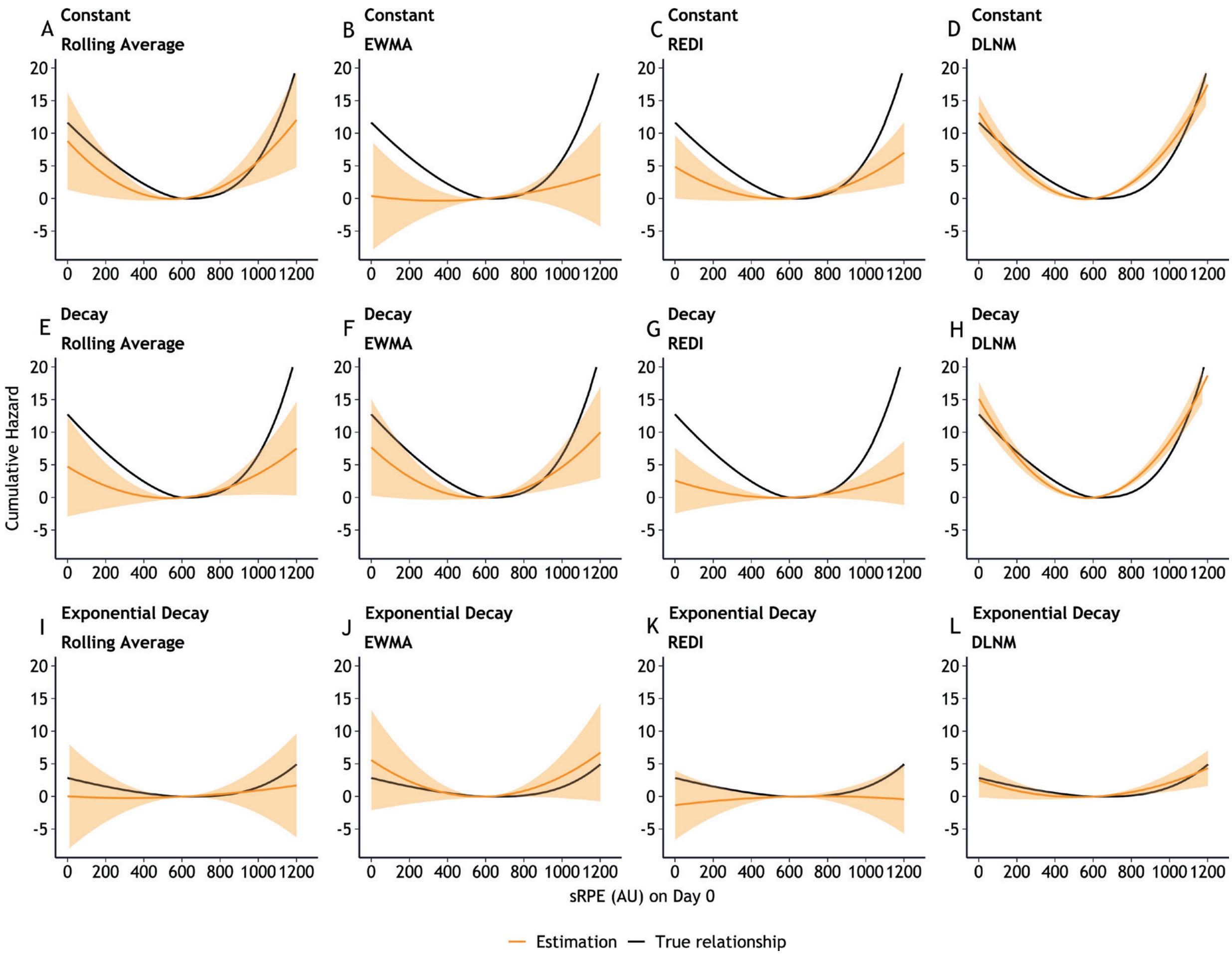

Figure 3 The relationship between absolute training load measured by the session rating of perceived exertion (sRPE) in arbitrary units (AUs) and the risk of injury on the current day (day 0) estimated by four different methods (yellow line), compared with the simulated, true relationship (black line). The y-axis denotes the cumulative hazard – the sum of all instantaneous risks of injury from the past up until the current day. Relationships were simulated under different scenarios, (A–D) constant: the risk of absolute training load is constant over time; (E–H) decay: the effect of absolute training load was at its highest on the current day (day 0) and reduced linearly for each lag day back in time; (I–L) exponential decay: the risk of absolute training load was at its highest on the current day (day 0) and reduced exponentially for each lag day back in time. Methods used to detect these effects were the rolling average, the exponential weighted moving average (EWMA), the robust exponential decreasing index (REDI), and the distributed lag non-linear model (DLNM). Yellow bands are 95% CIs. The figure shows one random simulation of 1900 performed.

scenarios (online supplemental figure S6), but for the Exponential Decay scenario, it had the lowest mean AIC and highest internal RMSE (table 1, online supplemental figure S6). Under all scenarios, DLNM had the lowest AIC in nearly 100% of simulations (online supplemental table S3). Although it was most frequently rank 1 internal RMSE for the Constant (52% of simulations) and Decay scenarios (57% of simulations), the rankings varied, and the Acute:Chronic Workload Ratio and Week-to-week %Δ were rank 1 ~23% of the time each (online supplemental table S3).

The Acute:Chronic Workload Ratio (ACWR) and week-to-week %Δ failed to discover a relationship between training load and injury under the Constant scenario (figure 4A, B). ACWR did not find a relationship under the Exponential Decay scenario, either (figure 4G). Both methods had wide confidence intervals, and ACWR fanned to higher uncertainty under higher levels of acute training load relative to chronic training load (figure 4). ACWR had marginally lower internal RMSE and lower AIC than week-to-week %Δ (table 1), and was rank 2 slightly more frequently than rank 3 (online supplemental table S3), except under the Exponential Decay scenario where the opposite was the case.

Handball example data analysis

The Distributed Lag Non-linear Model indicated, with high uncertainty, an increased risk of a health problem on the current day (HR (HR)>=1.2) for players with high internal load (sRPE above 4 000, figure 5A). This tapered to no effect if the training load was performed around a week ago (6 days before the current day, figure 5D), to a decreased risk of health problems the further in the past high training loads were sustained, to a HR of 0.75 on the

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

Table 1 Mean performance of methods used to estimate the effect of training load on injury risk (n simulations=1900).

Relationship	Method	External RMSE*	Internal RMSE	AIC	Coverage (%)	AW	Coverage MCSE
Absolute training load							
Constant	Rolling average	4.85	0.113547	1422.92	34.7	5.17478	0.90
	EWMA	4.77	0.113548	1423.42	36.3	5.17179	0.91
	REDI	5.53	0.113557	1424.10	20.3	3.40114	0.74
	DLNM	1.44	0.112434	1317.15	34.8	2.05600	0.95
Decay	Rolling average	5.38	0.113590	1421.80	30.2	5.16930	0.87
	EWMA	5.17	0.113587	1421.85	31.8	5.12554	0.88
	REDI	6.21	0.113605	1423.80	18.7	3.42154	0.71
	DLNM	1.55	0.112245	1295.30	32.4	2.07977	0.93
Exponential decay	Rolling average	2.13	0.113599	1424.65	85.0	5.54695	0.58
	EWMA	1.88	0.113588	1423.86	85.1	5.37141	0.61
	REDI	1.97	0.113603	1425.00	74.2	3.69208	0.64
	DLNM	0.76	0.113368	1407.08	81.6	2.02633	0.65
Relative training load (%Δ)†							
Constant	ACWR		0.113643	1426.16			
	Week-to-week %Δ		0.113646	1426.40			
	DLNM %Δ		0.113627	1389.28			
Decay	ACWR		0.113615	1424.73			
	Week-to-week %Δ		0.113617	1425.12			
	DLNM %Δ		0.113553	1383.52			
Exponential decay	ACWR		0.113565	1423.33			
	Week-to-week %Δ		0.113566	1423.27			
	DLNM %Δ		0.113700	1401.39			

*Monte Carlo SE for RMSE was <0.001 for all simulations. The scale of the RMSE depends on the scale of the coefficients, and it is therefore only interpretable by comparing values in the same analysis – the values cannot be interpreted in isolation.
†Due to differences in scale between methods and simulation for relative training load, external RMSE, coverage, and AW could not be calculated in a comparable manner.
ACWR, acute:chronic workload ratio; AIC, Akaike's information criterion; AW, average width of 95% CIs; Coverage, coverage of 95% CIs; DLNM, distributed lag non-linear mode; EWMA, exponentially weighted moving average; MCSE, Monte Carlo Standard Error; REDI, robust exponential decreasing index; RMSE, root-mean-squared error.

27th day before the current day (figure 5B). The cumulative risk was increased if an individual performed no training in the past and had high internal training load on the current day (figure 5C). None of the effects were significant (p>=0.8) and confidence intervals were broad (online supplemental table S4).

In the application of DLNM on a handball cohort, we were hampered by poor data quality. Also, due to anonymization, few covariates were available for confounder adjustment. The effects may have been spurious. We have included the analysis only as an illustration of how to use the DLNM in practice.

DISCUSSION

This is the first simulation study to explore methods for assessing the cumulative effect of long-term training load on injury or health problem risk in team sports. The Distributed Lag Non-linear Model (DLNM) had the highest combined accuracy and precision, the highest certainty, and the best model fit for almost all studied scenarios. It was the only method capable of exploring both the effects of the magnitude of training load and the time-dependent effects of past training load exposure.

Modelling methods for absolute training load

For determining the cumulative effect of the absolute training load, the Rolling Average was outclassed by the Exponentially Weighted Moving Average (EWMA). When the effect of absolute training load was simulated to be the same regardless of the distance in time since the current day – the scenario in which Rolling Average was thought to be appropriate – Rolling Average was only marginally better than the EWMA. EWMA had a better fit under the more realistic scenarios where the effects of training load decayed based on distance in time,

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

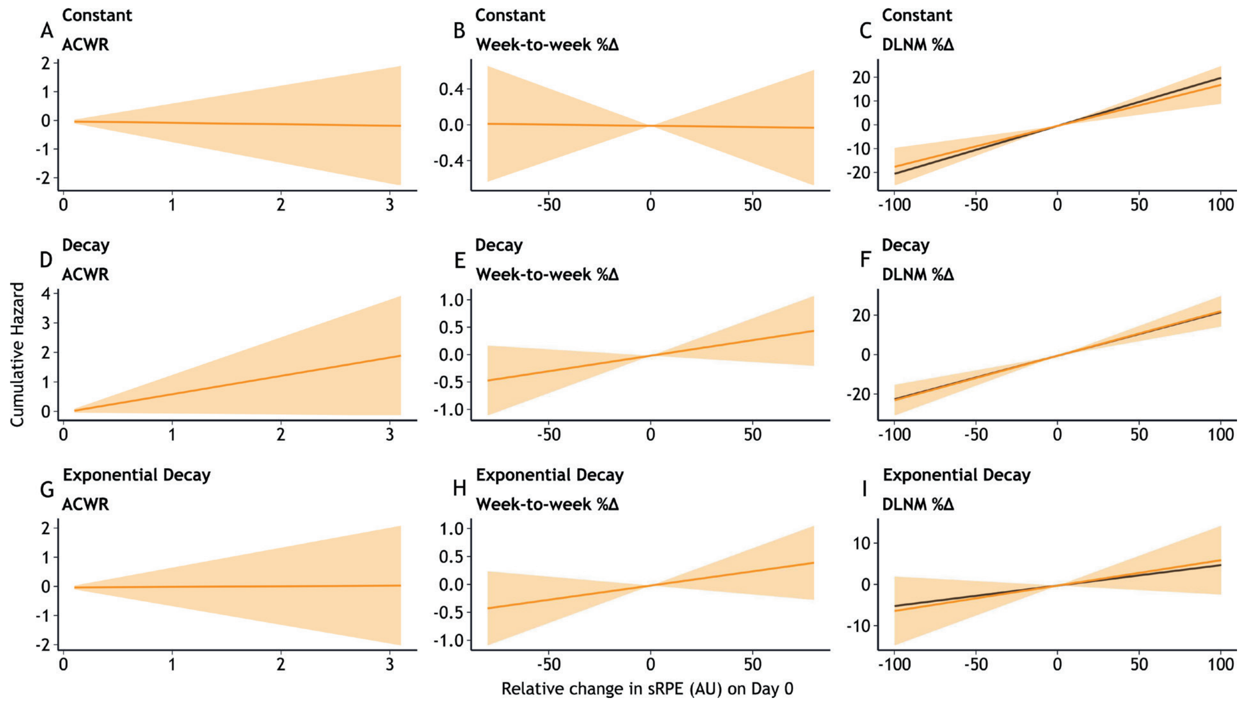

Figure 4 The relationship between relative training load measured in the daily percentage change of session rating of perceived exertion (sRPE) in arbitrary units (AUs) and the risk of injury on the current day (day 0) is estimated by three different methods (yellow line). The y-axis denotes the cumulative hazard – the sum of all instantaneous risks of injury from the past up until the current day. Relationships were simulated under different scenarios, (A–C) constant: the risk of relative training load was constant over time; (D–F) decay: the effect of relative training load was at its highest on the current day (day 0) and reduced linearly for each lag day back in time; (G–I) exponential decay: the risk of relative training load was at its highest on the current day (day 0) and reduced exponentially for each lag day back in time. Methods used to detect these effects were the acute:chronic workload ratio (ACWR), the week-to-week percentage change (%Δ) and the distributed lag non-linear model (DLNM) on daily percentage change Δ%. The DLNM, being on the same scale as the simulation, is also compared with the true, simulated relationship (black line). Yellow bands are 95% CIs. The figure shows one random simulation of 1900 performed.

both linearly and exponentially. This is in line with the concerns raised by Menaspà,[28] that the rolling average fails to take into account that training load performed in the past contributes less to injury risk than recent training load

The Robust Exponential Decreasing Index (REDI) was also outperformed by EWMA, under both scenarios where the training load effect decayed based on distance in time. Across the board, REDI had the highest RMSE, highest AIC, and lowest coverage of 95% confidence intervals. Although it had better RMSE under the Exponential Decay scenario than the rolling average, it erroneously estimated that higher internal training loads decreased injury risk (inverse relationship), when it was actually the opposite (ie, higher training load increased injury risk). REDI has previously been compared on observed training load values where the true relationship between training load and injury was unknown,[35] and it was recommended for its ability to handle missing data.[16] We believe that using imputation methods is more suitable for longitudinal data,[33] and in such cases, the advantage of specifying weights on missing observations is no longer applicable. REDI was among the methods that do not require a full time period (ie, 28 days) before the first calculation, but for comparability, we had to run it with the same limitation as the other methods. Arguably, it may therefore have performed better in a real study. On the other hand, this would also have been the case for the Distributed Lag Non-Linear Model (DLNM), which was vastly superior to all other methods analysed, even with this constraint.

DLNM had the lowest mean RMSE, AIC, and narrowest 95% CI width compared with the other three methods for all scenarios. The DLNM was the only method that did not require subjective aggregation. Aggregation distillates the information available in the data to a summary, and these summaries are all the Cox regression model must work with. This increases the uncertainty of the estimates. In contrast, DLNM uses all the information available in the data.[12] Furthermore, no subjective determination of time-lag weights is required. Using splines or fractional polynomials, it can explore non-linearity in both the magnitude of the effect of absolute training load and in the time-dependent effects.[15]

While it performed best compared with other methods, DLNM was unable to model the “Direct, then inverse” scenario. This scenario was built on the theory that training load exposure the current week increase risk

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

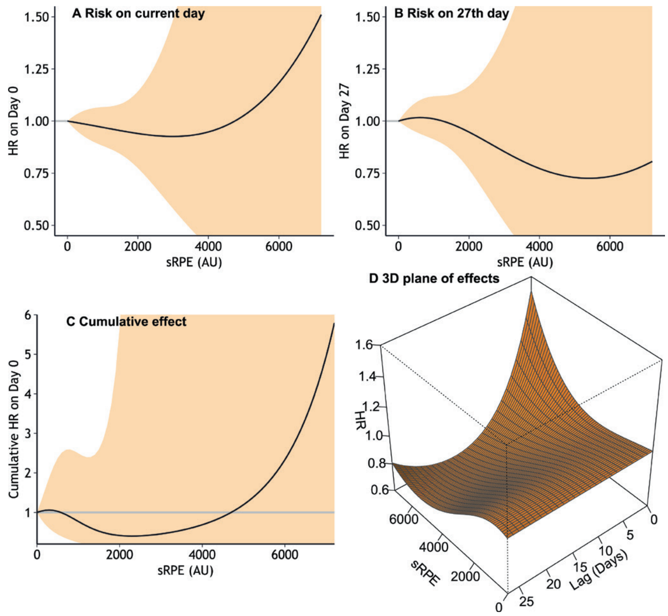

Figure 5 Explorations of the relationship between training load and the risk of suffering a health problem in a Norwegian elite youth handball cohort. Training load is measured by the session rating of perceived exertion (sRPE) in arbitrary units (AUs), shown on all x-axes. The health problem risk is measured by the Hazard Ratio (HR). HR >1 indicates an increased instantaneous health problem risk compared with an individual who had no training load (sRPE=0), <1 a decreased risk. Figure part A shows the risk of a health problem on the y-axis for each level of sRPE on the x-axis, given that the sRPE is sustained on the current day (day 0). Figure part B shows the same figure, given that the sRPE is sustained on the 27th lag day (4 weeks prior). Figure part C shows the cumulative HR – the collective risk of a health problem on the current day given the sRPE sustained in all the days prior to the current day. Finally, figure part D shows the risk relationship between absolute training load (sRPE) on the x-axis and the time since the training was sustained (lag) on the y-axis, where 0 is the current day and 27 is 4 weeks in the past. Risk in HR is on the z-axis. Yellow bands in (A–C) are the 95% CIs surrounding the estimates. The predictions pertain to a 17-year-old female. Based on 471 health problems from 205 handball players.

while those sustained the previous 3 weeks reduce risk.[8] Higher sample sizes than those in the current simulation may be needed to discover such a complex shape, if it were to exist. The splines may have required more than three knots, and linear splines may have been a better option than cubic splines to discover the sudden change in direction of effect.

Modelling methods for relative training load

Studying the relative training load proved challenging, as all methods compared were on different scales. According to the AIC, the most comparable metric,[12] DLNM had the best model fit under all scenarios. Given that we simulated an effect on the risk of injury based on the symmetrized percentage change from 1 day to the next, this was to be expected. The week-to-week percentage change and ACWR assume that day-to-day differences are of little to no importance. Currently, the time-period of relative training load that is relevant towards injury risk is debatable[36]; a calendar week may be arbitrary for many sports. We argue that if DLNM can detect the effect of day-to-day relative change, it should be flexible enough to detect less granular effects. In particular, team sports such as football often operate in micro-cycles of days since the previous match up to and including the next match.[15] However, it would still be up to the researcher to calculate percentage changes on time periods of their choosing before running DLNM, with the inherent difficulties of ratios.[25]

Even with the symmetrized percentage change, the percentage change cannot be calculated if the numerator or denominator is zero. Recovery days are an important aspect of training load history and must be evaluated to fully understand the effects of training load. This is a challenge that remains unsolved.

An application of distributed lag non-linear models in handball

The Distributed Lag Non-linear Model was able to explore non-linear time-dependent effects in the observed Norwegian youth elite handball data. The results had a high degree of uncertainty (p>=0.8), and we caution against considering them as evidence of a causal or associative relationship. They nevertheless illustrate how DLNM can be used in practice. DLNM can show how different levels of training load affects risk, and also how the effects changes with the distance in time since the training load exposure. It can also show the combined effect of these two dimensions and estimate the cumulative effect. However, performing DLNM and the corresponding visualisations in a training load and injury or health problem risk study may require collaboration with a statistician.[37] In addition, large sample sizes and good data quality may be needed to meet the complexity of the training load and injury risk relationship. In the handball data, 471 health problems occurred in 205 participants. As this was insufficient, future research may require even more participants for an accurate measure of effect.

Limitations

To feasibly analyse all results in a single article, we had to limit the number of methods compared in the simulations. This meant that two recently-proposed methods of relative training load were not among the compared methods.[38,39] Additionally, different variants of the ACWR were not considered, as these have been explored extensively in other studies.[30,40]

All methods in the simulation were run with the same specification for all scenarios to ensure consistency and comparability. In a real study, clinical rationale and hypothesis, as well as sensitivity analyses of model fit, would aid in determining the number and location of knots in splines for DLNM, the lambda value for EWMA

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

and REDI, and the time-periods for RA, EWMA, REDI and ACWR. Therefore, the flexibility of methods was not fully explored. In addition, for the relative training load, the simulation assumed that daily differences had an effect, an assumption that favoured DLNM, which has superior flexibility compared with the other methods. This advantage may be less prominent if stricter assumptions (ie, differences at the micro-cycle level) can be made[15]; however, we believe that the flexibility of the DLNM is one of its greatest strengths, rendering it useful in a wide range of situations.

CONCLUSION

The Distributed Lag Non-Linear Model is ideal for exploring the cumulative effect of the absolute training load and relative training load on injury risk, while accounting for time-dependent effects. For causal studies where training load is not the exposure of interest, but a confounder in need of adjustment, using the Exponentially Weighted Moving Average for the absolute training load is an alternative.

Twitter Lena Kristin Bache-Mathiesen @lena_kbm and Torstein Dalen-Lorentsen @torsteindalen

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

Contributors All authors conceptualized the study. Authors TEA, TD-L, and BC determined the aim and scope of the study, and the study design, with input from LKB-M and MWF. Author LKB-M performed simulations, statistical analyses, and wrote the manuscript under supervision from authors MWF and TEA. All authors contributed with notable critical appraisal of the text and approved the final version. LKB-M was the guarantor of this study.

Funding The Oslo Sports Trauma Research Centre provided all funding for performing this study. The Oslo Sports Trauma Research Center has been established at the Norwegian School of Sport Sciences through generous grants from the Royal Norwegian Ministry of Culture, the South-Eastern Norway Regional Health Authority, the International Olympic Committee, the Norwegian Olympic Committee & Confederation of Sport, and Norsk Tipping AS.

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 applicable.

Ethics approval This study involves human participants and was approved by The Norwegian Premier League football study and Norwegian elite youth handball study was approved by the Ethical Review Board of the Norwegian School of Sport Sciences, and by the Norwegian Centre for Research Data, with registration numbers 722773, and 407930, respectively. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available upon reasonable request. All data relevant to the study are included in the article or uploaded as supplementary information. All data relevant to the study are included in the article, are available as supplementary files, or available upon reasonable request. The anonymous training load variable from the Norwegian Premier League football data, and all statistical programming code, is available in a GitHub repository. The anonymised Norwegian elite youth handball data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

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 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 Med* 2019:1–20.
2 Soligard T, Schwellnus M, Alonso J-M, *et al*. How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury. *Br J Sports Med* 2016;50:1030–41.
3 Bittencourt NFN, Meeuwisse WH, Mendonça LD, *et al*. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition-narrative review and new concept. *Br J Sports Med* 2016;50:1309–14.
4 Stern BD, Hegedus EJ, Lai Y-C. State dependence: does a prior injury predict a future injury? *Phys Ther Sport* 2021;49:8–14.
5 Windt J, Gabbett TJ. How do training and competition workloads relate to injury? the workload-injury aetiology model. *Br J Sports Med* 2017;51:428–35.
6 Kalkhoven JT, Watsford ML, Coutts AJ, *et al*. Training load and injury: causal pathways and future directions. *Sports Med* 2021;51:1137–50.
7 Impellizzeri FM, Marcora SM, Coutts AJ. Internal and external training load: 15 years on. *Int J Sports Physiol Perform* 2019;14:270–3.
8 Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med* 2016;50:273–80.
9 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;53:61–8.
10 Bhaskaran K, Gasparrini A, Hajat S, *et al*. Time series regression studies in environmental epidemiology. *Int J Epidemiol* 2013;42:1187–95.
11 Richardson DB. Latency models for analyses of protracted exposures. *Epidemiology* 2009;20:395–9.
12 Gasparrini A. Modeling exposure-lag-response associations with distributed lag non-linear models. *Stat Med* 2014;33:881–99.
13 Williams S, West S, Cross MJ, *et al*. Better way to determine the acute:chronic workload ratio? *Br J Sports Med* 2017;51:209–10.
14 Nielsen RO, Bertelsen ML, Møller M, *et al*. Methods matter: exploring the 'too much, too soon' theory, part 1: causal questions in sports injury research. *Br J Sports Med* 2020;54:1119–22.
15 Bache-Mathiesen LK, Andersen TE, Dalen-Lorentsen T, *et al*. Not straightforward: modelling Non-linearity in training load and injury research. *BMJ Open Sport Exerc Med* 2021;7:e001119.
16 Moussa I, Leroy A, Sauliere G, *et al*. Robust exponential decreasing index (REDI): adaptive and robust method for computing cumulated workload. *BMJ Open Sport Exerc Med* 2019;5:e000573.
17 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.
18 Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Stat Med* 2019;38:2074–102.
19 Theron GF. *The use of data mining for predicting injuries in professional football players*. University of Oslo, 2020.
20 Team RC. *R: a language and environment for statistical computing*. Vienna, Austria: Foundation for Statistical Computing, 2021.
21 Sylvestre M-P, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Stat Med* 2009;28:3437–53.

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

22 Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *J Stat Softw* 2011;43:1.
23 Bache-Mathiesen LK. Modelling the cumulative effect of training load on injury risk: R code Repository, 2022. Available: https://github.com/lenakba/modelling-training-load
24 Foster C, Florhaug JA, Franklin J, *et al*. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15:115.
25 Curran-Everett D, Williams CL. Explorations in statistics: the analysis of change. *Adv Physiol Educ* 2015;39:49–54.
26 Dalen-Lorentsen T, Andersen TE, Bjørneboe J, *et al*. A cherry, ripe for Picking: the relationship between the Acute-Chronic workload ratio and health problems. *J Orthop Sports Phys Ther* 2021;51:162–73.
27 Wang A, Healy J, Hyett N, *et al*. A systematic review on methodological variation in acute:chronic workload research in elite male football players. *Sci Med Footb* 2021;5:18–34.
28 Menaspà P. Are rolling averages a good way to assess training load for injury prevention? *Br J Sports Med* 2017;51:618.1–9.
29 Ryan MR, Napier C, Greenwood D, *et al*. Comparison of different measures to monitor week-to-week changes in training load in high school runners. *Int J Sports Sci Coach* 2021;16:370–9.
30 Coyne JOC, Nimphius S, Newton RU, *et al*. Does mathematical coupling matter to the acute to chronic workload ratio? A case study from elite sport. *Inter J f Spo Physi Perfor* 2019;14:1447–54.
31 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.
32 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.
33 Bache-Mathiesen LK, Andersen TE, Clarsen B, *et al*. Handling and reporting missing data in training load and injury risk research. *Sci Med Foot* 2021;55:1–13.
34 Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014;48:1287–93.
35 Sedeaud A, De Larochelambert Q, Moussa I, *et al*. Does an Optimal Relationship Between Injury Risk and Workload Represented by the "Sweet Spot" Really Exist? An Example From Elite French Soccer Players and Pentathletes. *Front Physiol* 2020;11:1034.
36 West S, Williams S, Cazzola D. 119 The acute:chronic workload ratio: why one size does not fit all. *Br J Spor Med* 2021;55:A48.
37 Sainani KL, Borg DN, Caldwell AR, *et al*. Call to increase statistical collaboration in sports science, sport and exercise medicine and sports physiotherapy. *Br J Sports Med* 2021;55:118–22.
38 Afonso J, Nakamura FY, Canário-Lemos R, *et al*. A novel approach to training Monotony and Acute-Chronic workload index: a comparative study in soccer. *Front Sports Act Liv* 2021;3:149.
39 Tysoe A, Moore IS, Ranson C, *et al*. Bowling loads and injury risk in male first class county cricket: Is 'differential load' an alternative to the acute-to-chronic workload ratio? *J Sci Med Sport* 2020;23:569–73.
40 Wang C, Stokes T, Vargas JT, *et al*. Predicting Injury Risk Over Changes in Physical Activity in Children Using the Acute:Chronic Workload Ratio. *Am J Epidemiol* 2022;191:665–73.

BMJ Open Sport Exerc Med: first published as 10.1136/bmjsem-2022-001342 on 30 May 2022. Downloaded from http://bmjopensem.bmj.com/ on April 20, 2024 by guest. Protected by copyright.

Supplementary I: Results

FIGURES

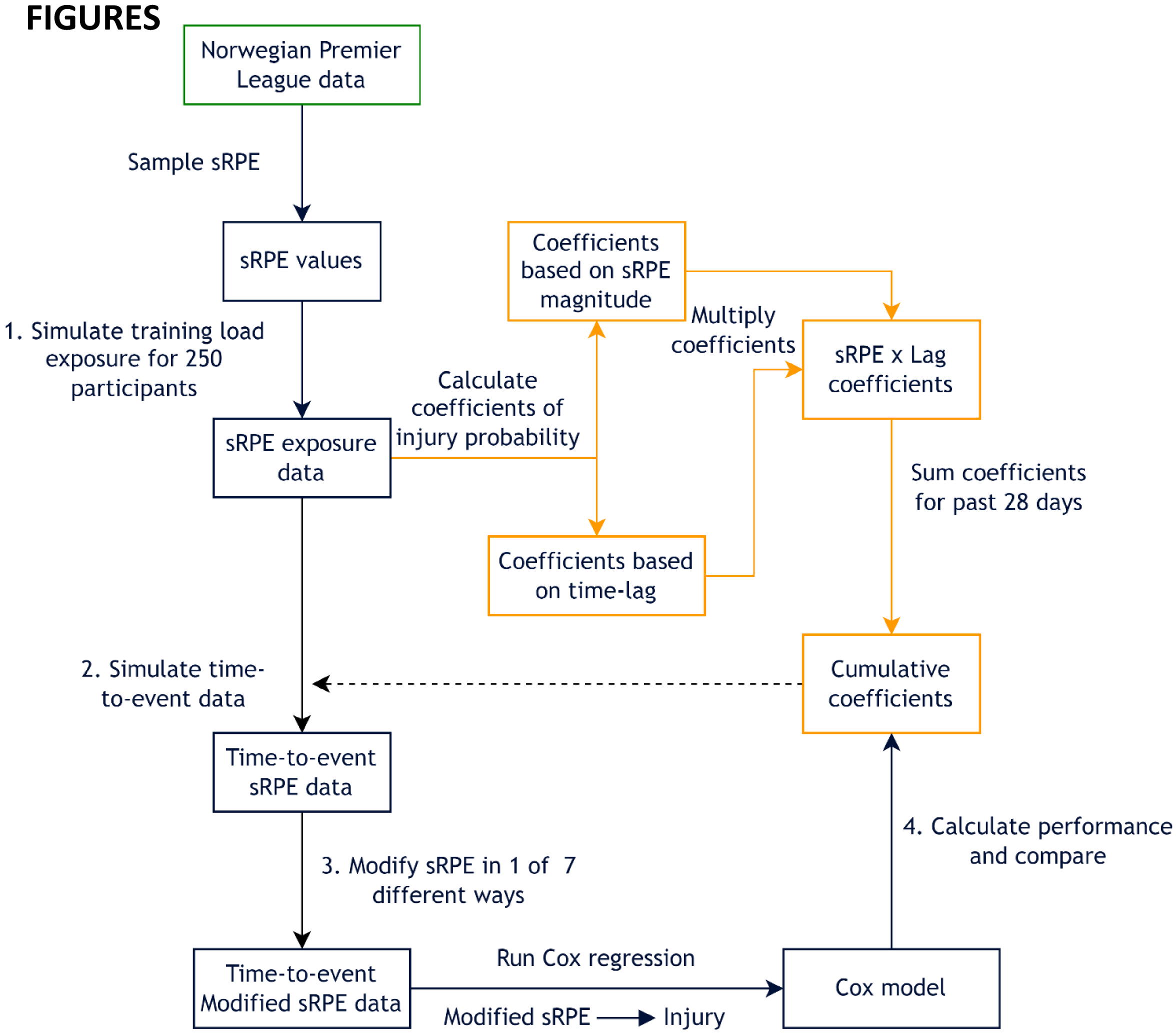

Figure S1. Summary of the simulation workflow. In Step 1, training load exposure measured by session Rating of Perceived Exertion (sRPE) was extracted from the Norwegian Premier League dataset and used to simulate training load exposure for 250 participants across 300 days. In Step 2, injury probabilities were calculated based on the cumulative training load observed the last 28 days; a combination of effect from both the magnitude of the training load (level of sRPE or %ΔsRPE) and the time since the training load occurred. Injuries were simulated based on these probabilities to generate time-to-event data. In Step 3, the absolute and relative training load exposures were modified and modelled in seven different Cox regression models. Finally, in Step 4, performance measures were calculated, and the accuracy of the different Cox models to detect the simulated relationship was assessed. Steps 1–4 were repeated 1 900 times for each of seven different simulated relationships (four for sRPE and three for %ΔsRPE) and each of seven methods.

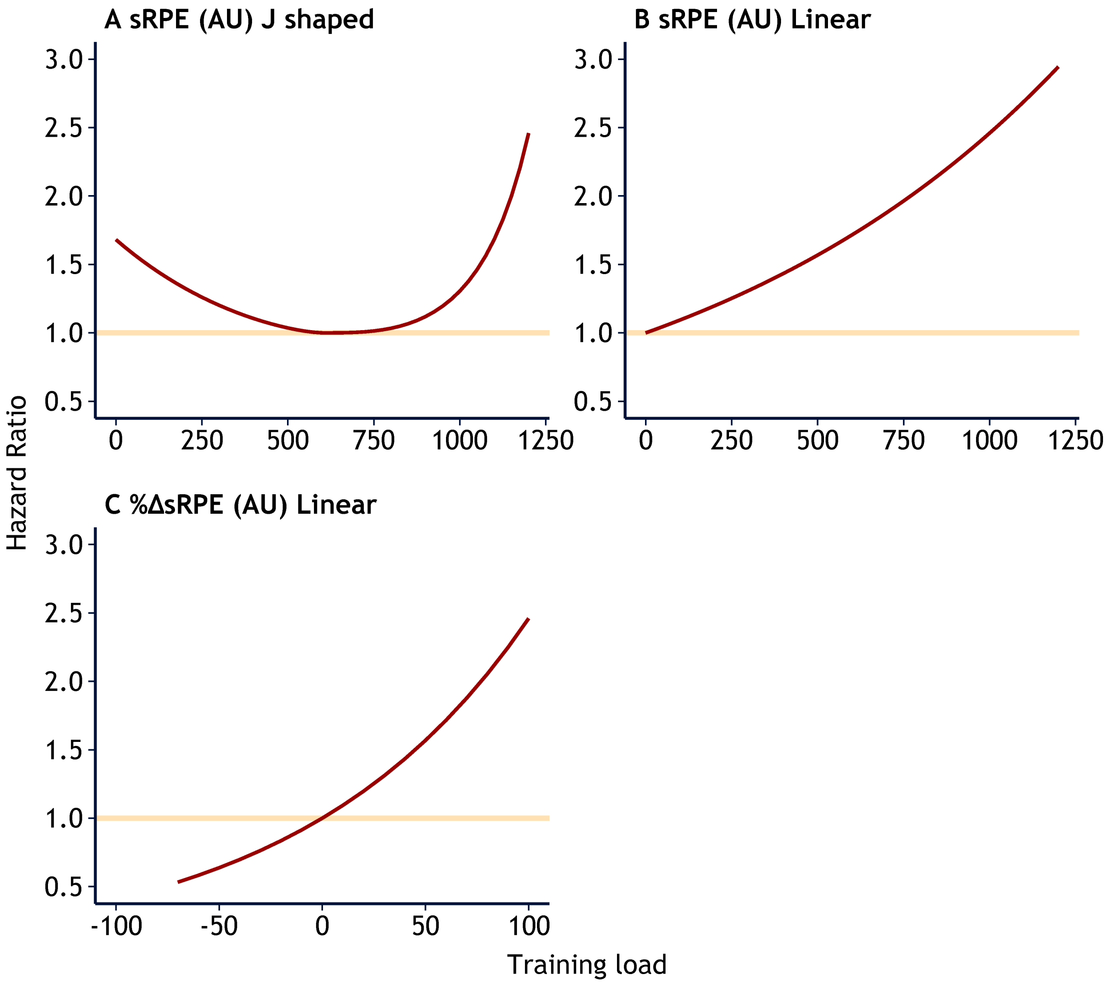

Figure S2. The simulated relationships between training load and injury risk, independent of the time since the training load exposure. Injury risk is measured by the Hazard Ratio (HR), where values > 1 (above the yellow line) indicates an increased risk and values < 1 (below the yellow line) indicates a decreased risk. Shown for (A–B) the absolute training load measured by the session Rating of Perceived Exertion (sRPE) measured in Arbitrary Units (AU), and (C) the relative training load compared to the previous day measured by the symmetrized percentage difference (%Δ) in sRPE. The absolute training load exposure was simulated with two different relationships, one J-shaped (A), and one linear (B).

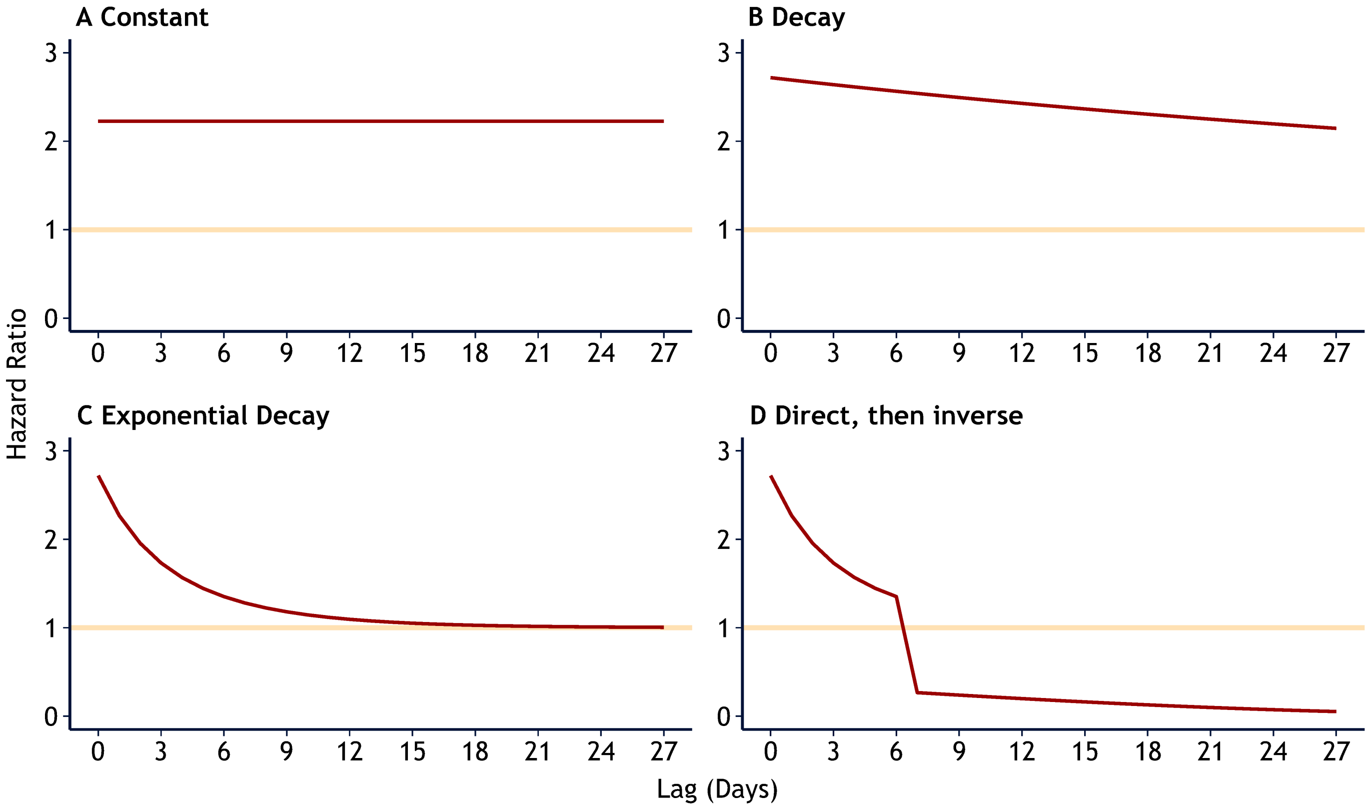

Figure S3. The simulated relationships between the time since current day (Day 0) that the training load exposure was sustained, and injury risk. Injury risk is measured by the Hazard Ratio (HR), where values > 1 (above the yellow line) indicates an increased risk and values < 1 (below the yellow line) indicates a decreased risk. The four risk shapes were (A) Constant, where the risk of training load is constant over time; (B) Decay, where the effect-size of the effect of training load is at its highest on the current day (Day 0) and is reduced for each lag day back in time; (C) Exponential Decay, where the risk of training load is at its highest on the current day (Day 0) and is reduced exponentially for each lag day back in time; (D) Direct, then inverse; where training load increases injury risk during the current week (Day 0–Day 6), but decreases injury risk thereafter. Training load had no effect after the 27^{th} lag day (4 weeks) in all four scenarios (not shown).

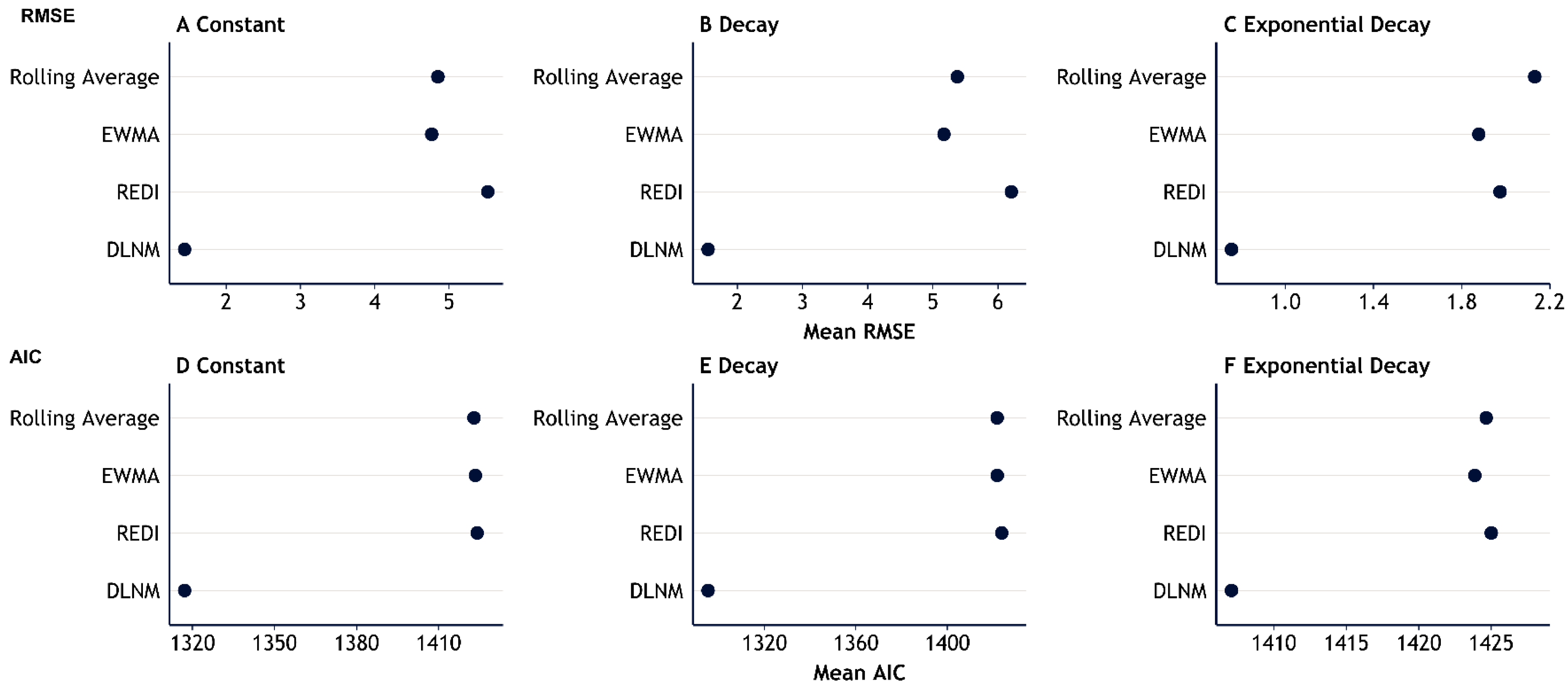

Figure S4. The mean Root-Mean-Squared Error (RMSE) and mean Akaike's Information Criterion (AIC) across 1 900 simulations of estimating the effect of absolute training load on injury risk. Due to variation in the effect sizes, calculations yield different scales for RMSE and AIC (x-axis) between relationship shapes; they cannot be compared between the three shapes, only within each shape. EWMA = Exponentially Weighted Moving Average; DLNM = Distributed Lag Non-Linear Model; REDI = Robust Exponential Decreasing Index. RMSE is calculated on the difference between the predicted risk and the simulated, true risk (External RMSE).

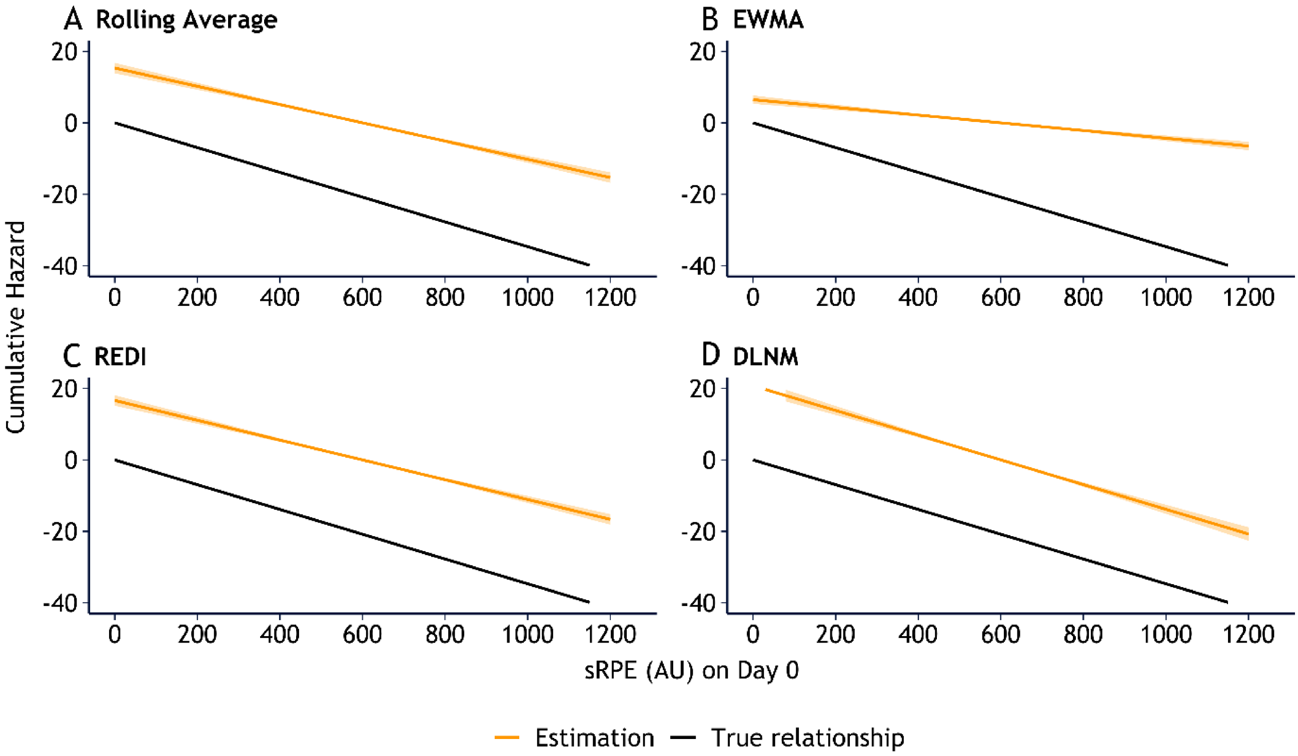

Figure S5. The relationship between absolute training load measured by the session Rating of Percieved Exertion (sRPE) in arbitrary units (AU) and the risk of injury on the current day (Day 0) estimated by four different methods (yellow line), compared with the simulated, true relationship (black line). The relationship scenario was “Direct, then inverse”, where training load increases injury risk during the current week (Day 0–Day 6), but decreases injury risk thereafter (Day 7–Day 27). The Y axis denotes the cumulative hazard – the sum of all instantaneous risks of injury from the past up until the current day. Methods used to detect these effects were (A) the Rolling Average, (B) the Exponential Weighted Moving Average (EWMA), (C) The Robust Exponential Decreasing Index (REDI), and (D) the Distributed Lag Non-Linear Model (DLNM). Yellow bands are 95% confidence intervals. The figure shows 1 random simulation of 1 900 performed.

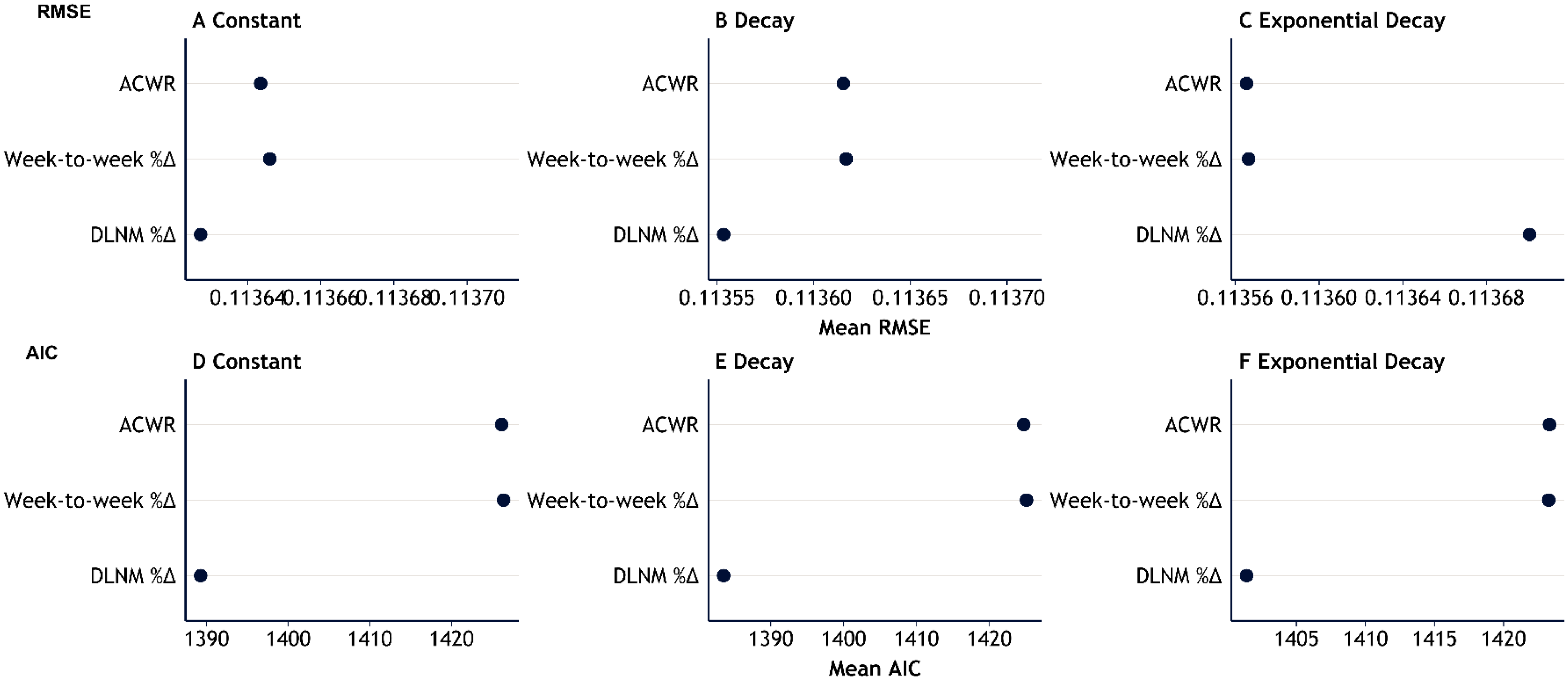

Figure S6. The mean Root-Mean-Squared Error (RMSE) and mean Akaike's Information Criterion (AIC) across 1 900 simulations of estimating the effect of relative training load on injury risk. Due to variation in the effect sizes, calculations yield different scales for RMSE and AIC (x-axis) between relationship shapes; they cannot be compared between the three shapes, only within each shape. ACWR = Acute:Chronic Workload Ratio; DLNM = Distributed Lag Non-Linear Model. RMSE is calculated on the model residuals (Internal RMSE).

TABLES

Table S1. The percentage of 1 900 simulations where methods of absolute training load had the lowest RMSE and AIC (Rank 1), had the 2nd lowest RMSE and AIC (Rank 2), and so on.

Metric	Lag scenario	Rank	Rolling Average (%)	EWMA (%)	REDI (%)	DLNM (%)
RMSE	Constant	1	2	1	0	97
		2	31	45	22	2
		3	52	27	21	1
		4	15	27	58	0
	Decay	1	1	1	0	98
		2	29	48	21	2
		3	54	26	19	0
		4	15	25	60	0
	Exponential Decay	1	11	13	13	63
		2	19	28	26	27
		3	36	27	29	8
		4	34	31	32	3
	Direct, then inverse	1	0	0	1	99
		2	0	0	99	1
		3	100	0	0	0
		4	0	100	0	0
AIC	Constant	1	0	0	0	100
		2	31	39	31	0
		3	58	24	18	0
		4	11	38	51	0
	Decay	1	0	0	0	100
		2	31	45	24	0
		3	59	24	17	0
		4	10	31	59	0
	Exponential Decay	1	1	1	1	97
		2	19	52	28	2
		3	55	22	23	0
		4	26	25	48	1
	Direct, then inverse	1	0	0	0	100
		2	0	0	100	0
		3	100	0	0	0
		4	0	100	0	0

Abbreviations: AIC = Akaike's Information Criterion; EWMA = Exponentially Weighted Moving Average; DLNM = Distributed Lag Non-Linear Model; REDI = Robust Exponential Decreasing Index; RMSE = Root-Mean-Squared Error

Table S2. Mean performance of methods used to estimate the effect of absolute training load on injury risk under the "Direct, then inverse" scenario.

	Rolling Average	EWMA	REDI	DLNM
External RMSE[1]	21.1	22.6	20.9	20.8
Internal RMSE	0.111	0.113	0.106	0.101
AIC	1116	1373	910	790
Coverage[1]	0%	0%	0%	0%
AW	1.48	1.25	1.56	1.94

Abbreviations: AIC = Akaike's Information Criterion; AW = Average Width of 95% confidence intervals; Coverage = Coverage of 95% confidence intervals; EWMA = Exponentially Weighted Moving Average; DLNM = Distributed Lag Non-Linear Model; REDI = Robust Exponential Decreasing Index; RMSE = Root-Mean-Squared Error

[1] Monte Carlo Standard Error was < 0.001 for RMSE, and 0.5 for coverage of 95% confidence intervals for all methods.

Table S3. The percentage of 1 900 simulations where methods of relative training load had the lowest RMSE and AIC (Rank 1), had the 2nd lowest RMSE and AIC (Rank 2), and so on.

Metric	**Lag scenario**	**Rank**	**ACWR (%)**	**Week-to-week %Δ (%)**	**DLNM %Δ (%)**
RMSE	Constant	1	25	23	52
		2	49	49	2
		3	26	29	46
	Decay	1	23	21	57
		2	50	48	2
		3	28	31	41
	Exponential Decay	1	31	29	41
		2	48	50	2
		3	22	21	57
AIC	Constant	1	0	0	100
		2	56	44	0
		3	44	56	0
	Decay	1	0	0	100
		2	59	41	0
		3	41	59	0
	Exponential Decay	1	1	1	99
		2	49	51	0.5
		3	52	49	0.9

Abbreviations: ACWR = Acute:Chronic Workload Ratio; AIC = Akaike's Information Criterion; DLNM = Distributed Lag Non-Linear Model; RMSE = Root-Mean-Squared Error

Table S4. The model coefficients from a Cox regression estimating the relationship between training load and risk of injury in a handball cohort (n players = 205, n injuries = 472).

Term[1,2]		HR	95% CI Lower–Upper	SE	DF	p-value
sRPE 1		0.80	0.11–5.70	0.897	11.758	0.81
sRPE 2		0.99	0.87–1.13	0.059	11.909	0.88
sRPE 3		0.77	0.01–99.10	2.259	13.435	0.91
sRPE 4		0.96	0.70–1.33	0.150	13.445	0.81
Age		0.97	0.79–1.21	0.109	456.684	0.80
Sex						
	Female (Reference)	-	-	-	-	-
	Male	1.13	0.781-1.641	0.189	462.46	0.51

Abbreviations: CI = Confidence Interval; df = Degrees of Freedom; HR = Hazard Ratio; SE = Standard Error; sRPE = session Rating of Perceived Exertion

[1]The frailty term for within-individual variance was significant at p < 0.00001

[2]The sRPE terms are the four intervals demarcated by 3 knots in the restricted cubic splines

Supplementary II: Methods

FOOTBALL DATA SIMULATION

As recommended in O'Kelly, et al. [1], a study protocol was developed before initiation of simulations and analyses. Our methodology was focused on a causal research setting; however, the methods may also be applied in predictive research.[2] Simulation steps 1–4 detailed below are illustrated in online supplemental file 1 figure S1.

Step 1 Preparing data

We constructed different relationships between training load and injury based on a dataset of Norwegian Premier League male football players followed for 323 days (n = 36, mean age 26 years [Standard Deviation 4]). Training load was measured daily with the session Rating of Perceived Exertion (sRPE)[3]: the duration of the activity in minutes multiplied by the player's perceived intensity of the activity on a scale from 0 to 10. The players reported intensity and duration after completion of each training session or match,[4] using a mobile application (Athlete Monitoring, Moncton, Canada). The mean answering time was 0.01 days (SD = 0.2); 99% of prompts were answered within the same day, and the longest answering time was 4 days. Of 4 871 prompts, 650 (13%) Rating of Perceived Exertion observations were missing.[5] The relative training load from one day to the next was calculated with the symmetrized percentage change (%ΔsRPE).[6]

The most common study design in training load and injury risk studies is one team of athletes followed for one season.[7] By rough estimate, a football team suffers on average 40 injuries per team per season, not counting recurrent injuries.[8] The association between training load and injury is likely to be small to moderate,[9] therefore, one team followed for one season is unlikely of sufficient power to detect a relationship accurately,[10] and in most cases, studies will focus on a particular injury type, i.e. hamstring injury. We therefore simulated a medium-to-large-sized study: 250 participants (10 football teams), followed for a season (300 days).

Step 2 Simulating time-to-event data

We simulated injuries under different relationship scenarios with the sampled training load. For simplicity, only one injury was simulated per individual. This scenario may be unrealistic, as sports injuries may be sustained multiple times.[11] The methods for modelling training load considered in this study can, however, also be used with more complex statistical models for recurrent events.[12] The risk of injury at any given time was predetermined with a time-to-event Cox regression model with one covariate:

$$h(t) = h_0(t) * \exp(\beta x) \qquad \textit{Eq. 1}$$

Where h_0 is the baseline hazard, and $h(t)$ is the hazard at timepoint t. The timepoint at which an individual could be censored was drawn at random from a uniform distribution ranging from 0 to 600. Here, x represents the absolute training load, but it can be replaced with the relative training load, $\%\Delta x$. The coefficient β was the result of a bidimensional function on both the magnitude of the training load x, and the distance in time, the time lag l, from the timepoint t. We can write this more accurately:

$$h(t) = h_0(t) * \exp(s(x_t, \dots, x_{t-l}, \dots, x_{t-L})) \qquad \text{Eq. 2}$$

Here, the function s describes the relationship between training load x and the hazard of injury, measured over the lag interval $l = 0, \dots, L$ where L is the maximum lag. We denoted $l = 0$ to be the current day (Day 0), and the max lag was set at $L = 27$. This corresponds to 28 days (4 weeks).

The s function, $s(x_t, \dots, x_{t-L})$, can be defined in multiple ways.[13] We simulated s to be the cumulative sum of both a function on the magnitude of training load, the variable function $f(x)$, and a function on the distance in time from the current day, the lag function $w(l)$. This can be represented by:

$$s(x_t, \dots, x_{t-L}) = \sum_{l=0}^{L} f(x) \cdot w(l) \qquad \text{Eq. 3}$$

The shape of the relationship between the absolute training load and injury risk was simulated to be J-shaped (online supplemental file 1 figure S2A).[14] Under this assumption, the lowest point of risk was intermediate levels of training load. The highest was under high levels of training load. The variable function $f(x)$ was:

$$f(x) = \begin{cases} ((600 - x)/200)^\wedge 1.5/10, & x < 600 \\ ((x - 600)/200)^\wedge 3/30), & x \geq 600 \end{cases}$$

Where x was measured with the sRPE. For the relative training load, we simulated a linear relationship with injury risk (figure S2C). Higher loads on the current day compared to load on the previous day increases risk, whilst lower loads on the current day compared with the previous day reduces risk[15]:

$$f(\%\Delta x) = 0.009 * \%\Delta x$$

Here, $\%\Delta x$ was the symmetrized percent change from the previous day, ranging from -100% to 100%.

To compare methods ability to discover different time-dependent effects, the lag function $w(l)$ was defined in four different scenarios.

Constant. Across 4 weeks, the effect of training load has a constant effect each day (online supplemental file 1 figure S3A). Thereafter, training load has no effect. This was an overly simplistic base scenario.

$$w(l) = 0.8$$

Decay. Across 4 weeks, the effect of training load gradually decays for each day (figure S3B).[16] Thereafter, training load has no effect. This was hypothesized as a likely scenario if past training load has a direct effect on injury risk.

$$w(l) = \exp\left(-\frac{l}{100}\right)$$

Exponential decay. On the current day, training load has the highest risk of injury. The effect of training load drops exponentially the past 4 weeks (figure S3C). Thereafter, training load has no effect. This was hypothesized as a likely scenario if past training load has an indirect effect on injury risk.

$$w(\,l) = \exp\left(-\frac{l}{10}\right)^2$$

Direct, then inverse. Training load values on the current week (acute) increases risk of injury, whilst the training load values three weeks before the current week (chronic) decreases risk of injury (figure S3D)[17] Thereafter, training load has no effect. This hypothesis has recently been challenged.[18 19] Nevertheless, to ensure that modelling methods can uncover this relationship should it be true, we opted to include it regardless. The theory depends on chronic load amount as a surrogate measure for fitness, and acute load amount a surrogate measure for fatigue.[15] High loads relative to the previous time period are thought to increase risk, while low loads relative to the previous time period decrease risk: a linear relationship.[15 20 21] Therefore, for this time-lag scenario, we simulated a linear relationship with the absolute training load, and the relative load was not considered,

$$w(\,l) = \begin{cases} \exp\left(-\frac{l}{10}\right)^2, & l \le 6 \\ -\exp\left(\frac{l}{50}\right)^2, & l > 6 \end{cases}$$

The relationships constant, decay and exponential decay were used both for the absolute training load and for the relative training load. The “Direct, then inverse” relationship was only simulated for the absolute training load exposure. For this time-lag scenario, and for this time-lag scenario only, we simulated a linear relationship with the absolute training load (online supplemental file 1 figure S2B):

$$f(x) = 0.0009 * x$$

In summary, seven different relationships between training load and injury risk were simulated (figure 1–2). In a pilot of 100 simulations for each of the seven scenarios, the mean number of simulated injuries for 25 participants (a football team) was 18.7 per season; reasonably realistic of a small-to-moderate effect between training load and a specific injury type (i.e. a study on hamstring injury).

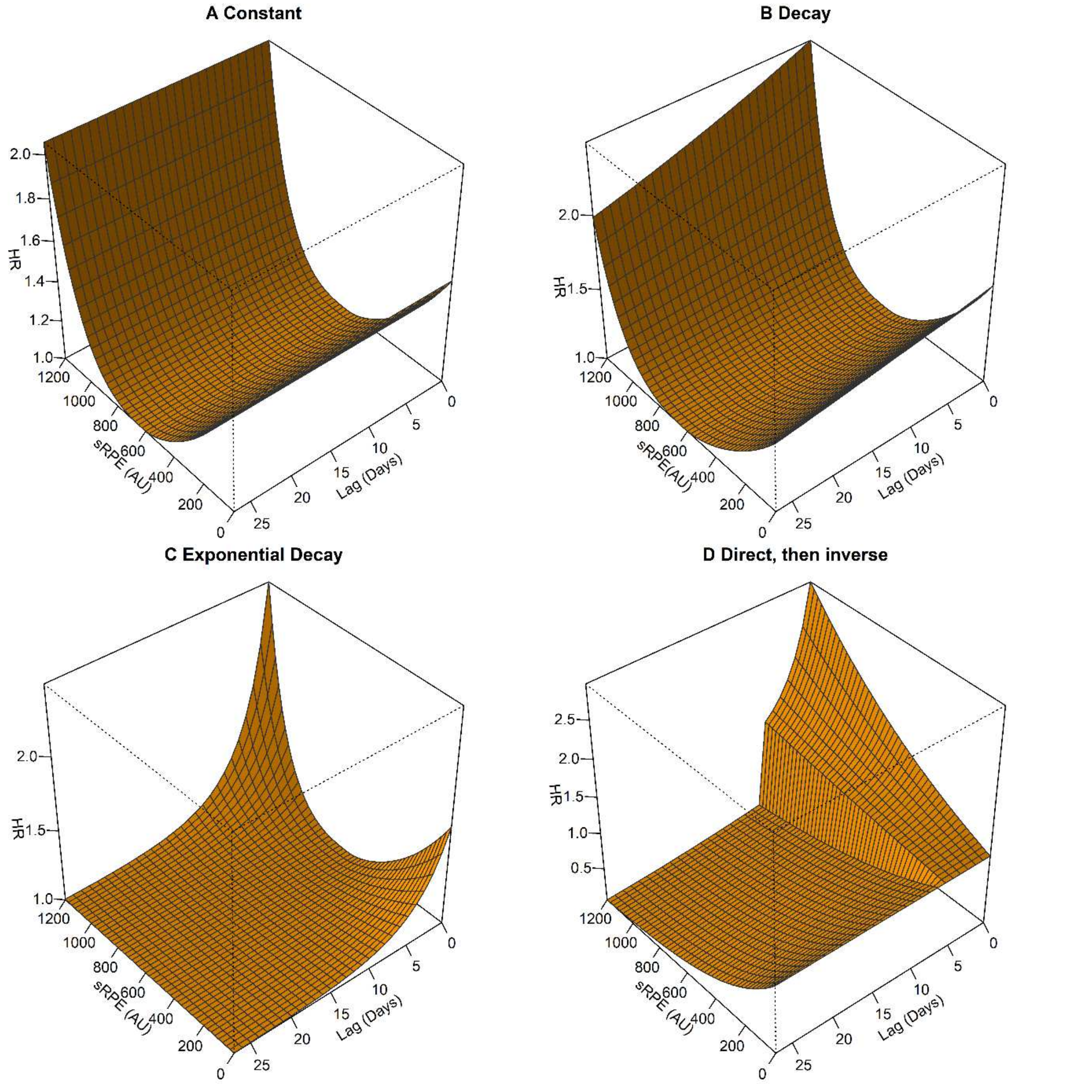

Figure 1. The four simulated relationships between absolute training load and injury risk. The relationships are a combination of the J-shaped function on the absolute training load exposure (online supplemental file 1 figure S2A) and the different functions on the time since training load was sustained (figure S3). Training load is measured with the session Rating of Perceived Exertion (sRPE), shown on the X-axis. The time since the current day (Day 0) is shown on the Y-axis, where 0 is the current day and 27 is the 27th day before the current day. On the Z-axis, the risk of injury is measured with the Hazard Ratio (HR), where HR > 1 indicates an increased risk, and HR < 1 indicates a decreased risk. The four risk shapes are (A) Constant, where the J-shaped risk of training load is constant over time; (B) Decay, where the effect-size of the J-shaped effect of training load is at its highest on the current day (Day 0) and is reduced linearly for each lag day back in time; (C) Exponential Decay, where the J-shaped risk of training load is at its highest on the current day (Day 0) and is reduced exponentially for each lag day back in time; (D) Direct, then inverse; where training load linearly increases injury risk during the current week (Day 0–Day 6), but linearly decreases injury risk thereafter. This was the shape simulated with a linear model on the absolute training load (figure S2B). Training load had no effect after the 27th lag day (4 weeks) in all four scenarios (not shown).

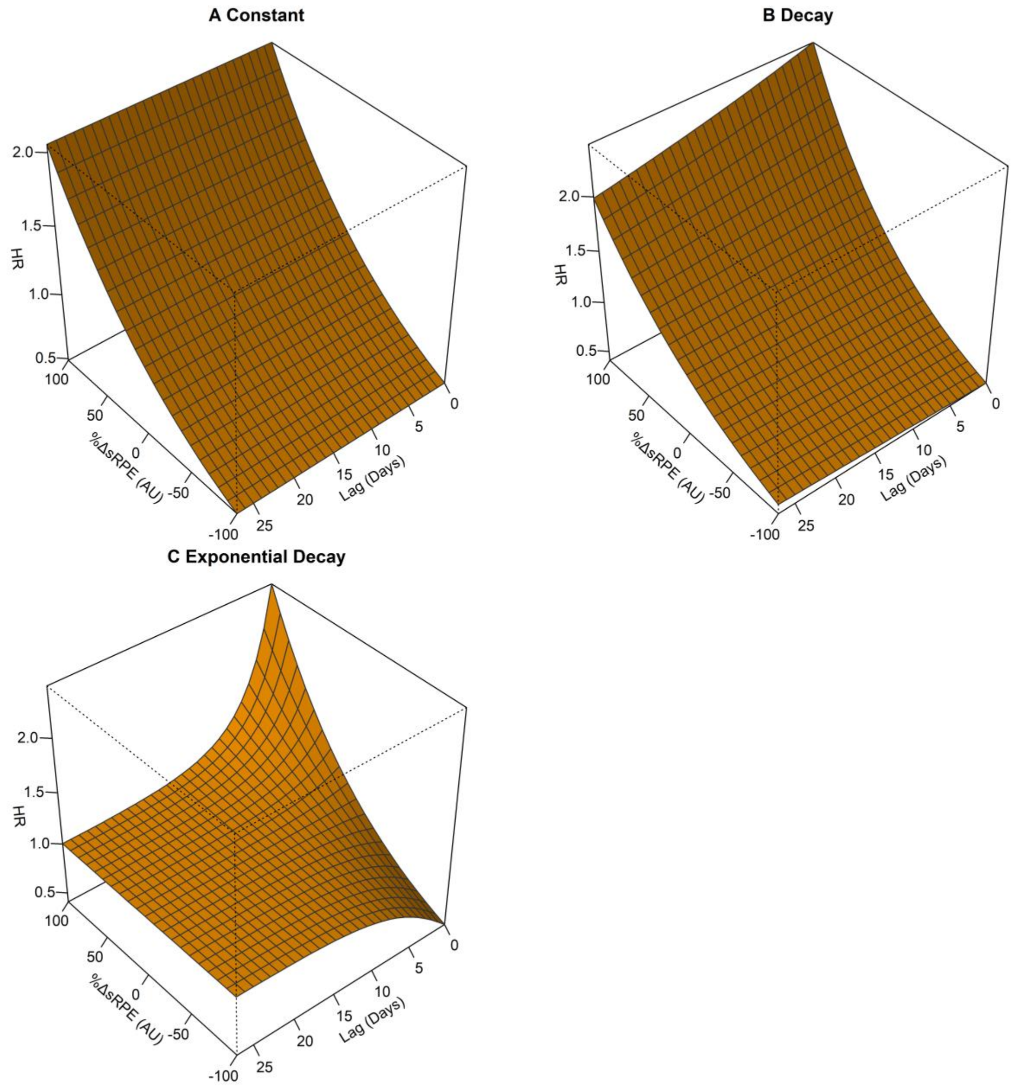

Figure 2. The three simulated relationships between relative training load and injury risk. The relationships are a combination of the linear function on the relative training load exposure (online supplemental file 1 figure S2C) and the different functions on the time since training load was sustained (figure S3). Relative training load is measured with the symmetrized percentage change (%Δ) in session Rating of Perceived Exertion (sRPE), shown on the X-axis. The time since the current day (Day 0), the number of lag days is shown on the Y-axis, where 0 is the current day and 27 is the 27th day before the current day. On the Z-axis, the risk of injury is measured with the Hazard Ratio (HR), where HR > 1 indicates an increased risk, and HR < 1 indicates a decreased risk. The four risk shapes are (A) Constant, where the linear risk of relative training load is constant over time; (B) Decay, where the effect size of the linear effect of relative training load is at its highest on the current day (Day 0) and is reduced linearly for each lag day back in time; (C) Exponential Decay, where the linear risk of training load is at its highest on the current day (Day 0) and is reduced exponentially for each lag day back in time. Training load had no effect after the 27th lag day (4 weeks) in all three scenarios (not shown).

Step 3 Modelling the time-dependent effect of training load on injury risk

Different methods of modelling training load were compared in their ability to uncover the seven predetermined relationships between training load and injury risk. A Cox regression model (Eq. 1) was used to estimate the relative risk of injury, where training load, x or $\%\Delta x$, was modified or modelled in three different ways for the absolute training load, and three different ways for the relative training load.

We chose the most frequently used methods in training load and injury research,[22-24] methods proposed as potential alternatives,[16 25] and a method developed to handle similar challenges in epidemiology.[26 27]

In the Cox regression model, regardless of method used to modify the absolute training load, the training load was modelled with a quadratic term under all time-lag scenarios except for the "Direct, then inverse", where a linear term was used. This was done to ensure methods were compared under the same conditions. Here, we assumed that a given researcher would have performed a sensitivity analysis before-hand to determine the need for a linear vs. non-linear shape.

A linear relationship was assumed between relative training load and injury risk, regardless of method used to modify the training load.

Absolute training load

Rolling average

Despite past critiques,[28] the rolling average (RA)[29] was the most frequently used method to account for the cumulative effects of training load in recent reviews.[23 30] Training load and injury risk studies that employ more advocated methods[16] still calculate the RA alongside the other calculations.[31-34] We therefore included this method in our comparison. For training load denoted x, the moving average RA is defined by:

$$RA = \frac{x_{k-n+1} + x_{k-n+2} + \cdots + x_k}{n}$$

Where n is the size of the time-lag window, in this study, 28 days. k denotes the last value in the time-lag window for an individual. For the first window, k = 28, for the second window, $k = 29$, and so on, up until the final window, k = 300. For each window, the first value is removed from the calculation, and the next value is added. For example, the first rolling average calculation is:

$$RA_1 = \frac{x_1 + x_2 + \cdots + x_{28}}{28}$$

The second rolling average calculation is:

$$RA_2 = \frac{x_2 + x_3 + \cdots + x_{29}}{28}$$

This sliding window of calculation can thus be generalized to:

$$RA_{today} = RA_{yesterday} + \frac{1}{n}(x_{k+1} - x_{k-L+1})$$

The method is intuitive and simple to calculate. An advantage is that it can be calculated on incomplete time-windows, given that n is defined as the number of training load values in the time sequence so far. For comparability with other methods, however, we calculated RA only from the 28th value and so on. The disadvantage is that rolling averages assume that training loads further back in time, and more recent training loads, contribute equally to injury risk.[16] The method provides no flexibility in the size or direction of effect for different time-lags.[35]

Exponentially weighted moving average

The exponentially weighted moving average (EWMA) is an extension of the rolling average. It accounts for the assumption that training load values further back in time contribute less to injury risk than training loads closer in time to the current day.[16] It has been recommended as an improvement over the rolling average, [16 36] and has been used in training load and injury risk studies since.[24 30 33] For training load denoted x, EWMA is:

$$EWMA_{today} = x_{today} + \lambda + ((1 - \lambda) + EWMA_{yesterday})$$

Where λ represents the decrease in effect depending on distance in time, by number of days n, up to a maximum of $n = 28$:

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

This choice of lambda is the same as in Williams, et al. [16] and Moussa, et al. [25].

A disadvantage of the EWMA is that a full window (28 days) must be completed before the calculation of the first EWMA. Any injuries sustained in this period are therefore not included in the analysis of injury risk. In addition, EWMA is constrained to an exponential weight only, and it cannot be calculated in the presence of missing values.[25]

Robust exponential decreasing index

The Robust Exponential Decreasing Index (REDI) has recently been proposed as an alternative over the EWMA,[25] and had improved performance in a training load and injury risk study.[37] For the lag interval $l = 0, \ldots, L$ where $l = 0$ is the current day, and L is the maximum lag 27, we can determine a vector of coefficients for each lag. Then, multiply the coefficients with the training load at each lag and sum these weighted training load values.

$$\text{Weighted x} = \sum_{l=0}^{\text{L}} \alpha_l^\lambda * \text{x}_\text{l}$$

The coefficient, α_l^λ is determined as follows:

$$\alpha_l^\lambda = \begin{cases} 0 & if\ x\ is\ missing \\ \exp(-\lambda * l) & if\ x\ is\ not\ missing \end{cases}$$

The λ weight has to be specified by the user, same as the EWMA method. The weighted training load values are then divided by the sum of the weights:

$$REDI = \frac{\text{Weighted x}}{\sum_{l=0}^{L} \alpha_l^\lambda}$$

The lower the lambda ($\lambda \to 0$), the greater the impact from past training load values. We chose lambda = 0.1 as it was the highest lambda value where training load on the 27th lag day still contributed to the cumulative effect.[25] Coincidentally, it was the same as used in Moussa, et al. [25], and is closest in behavior to the EWMA.

REDI is robust to missing data in training load, and like the rolling average, it can be calculated on incomplete time-windows. In addition, it may be more flexible than the EWMA in that the choice of lambda can fine-tune the weights to a specific sport or setting.[25]

Distributed lag non-linear model

In environmental epidemiology, modelling long-term effects – such as pollution or radon-exposure – is a common challenge. Although not entirely applicable to the challenges with training load, they do share the complexities of being long-term, weak-to-moderate protracted time-varying effects.

To recap, the relative risk of injury is considered to be the combined result of 1) the magnitude of exposure to training load, known as the exposure-response relationship, and 2) the distance in time from the current day (Day 0), the lag-response relationship.

To handle such effects, Bhaskaran, et al. [26] suggested using a so-called distributed lag model, a method initially developed in econometrics[38] and later applied to epidemiology.[39]

With Eq. 2, we explained how the β-coefficient for training load can be a result of the s function, $s(x_t, \ldots, x_{t-L})$. In a distributed lag model, the effects from the lag-response relationship is modelled with the lag-response function $w(l)$:

$$s(x_t, \ldots, x_{t-L}) = \sum_{l=0}^{L} x_{t-l} w(l)$$

When $w(l)$ is a constant function, this is equivalent to the rolling average.[13] Distributed lag models has been implemented in environmental epidemiology to handle cumulative, time-dependent effects.[26 40] The downside is the data-driven exploration of cut-offs,[35] and the assumption of a linear relationship between exposure, lag and response.[26]

To account for these issues, Bhaskaran, et al. [26] recommended using polynomial or splines to explore the long-term pattern in so-called Distributed Lag Non-linear Models (DLNM).

This has been applied to time-to-event data in medicine.[41 42] DLNMs allow non-linear modelling of the combined effect of the exposure-response and the lag-response relationships: the exposure-lag-response relationship.[27] The function s can be defined by crossing the variable function $f(x)$ and the lag function $w(x, l)$ and thus produce a bi-dimensional exposure-lag-response function $f(x) \cdot w(x, l)$:

$$s(x_t, \dots, x_{t-L}) = \sum_{l=0}^{L} f(x) \cdot w(x_{t-l}, l)$$

The exposure-response function $f(x)$, the function on the absolute training load, must be specified by the user. In the Cox regression model, $f(x)$ was modelled with a quadratic term, except for the "Direct, then inverse" time-lag scenario, where a linear term was used instead; same as for the other methods. The lag-response function $w(x, l)$ is the function for the time-dependent effect, and must also be specified by the user. Here, it was modelled with restricted cubic splines using 3 knots under all scenarios, since splines can explore non-linear shapes.[14] For a gentle introduction to DLNMs, see Gasparrini [13]. For more extensive mathematical exploration, see Gasparrini [27].

DLNM is a method which models, rather than modifies, training load. Therefore, no discarding of data, choice of time-blocks, or aggregation of training load values is necessary, and so, all information in the raw data is retained. Another advantage is that DLNM is flexible in the modelling of the exposure-response and the lag-response functions, both of which may be modelled with polynomials or splines at the user's discretion. This allows the exploration of non-linear and complex time-lag effects. On the other hand, modelling complex time-lag effects may require larger sample sizes, and model specification requires subjective choice.[13]

Relative training load

Week-to-week percentage change

In training load studies, it is common to divide the data into blocks of time.[43 44] The weekly sRPE is calculated by summing the daily sRPEs.[34] The percentage difference can then be calculated on the difference in sRPE between the current week and the previous week.[45 46] We included this method in the comparison as the most basic method of calculating relative training load. The percentage difference has a few disadvantages,[6] one being that it cannot be calculated when the denominator is zero. We therefore opted for the symmetrized percentage change, which has improved mathematical properties.[6] This calculation can be represented by:

$$\%\Delta \mathrm{W} = \frac{W_k - W_{k-1}}{W_k + W_{k-1}} * 100$$

Where k is the current week. In the same manner as the moving average, the week-to-week percentage change calculation moves iteratively from one week to the next.

The week-to-week percentage change is simple to calculate. Any injuries suffered in the first six days must be discarded before calculation of the first percentage difference. However, this is a small amount of data compared to some of the other methods compared. The main disadvantage is that it does not consider training load values further back in time than the previous week, and the time-block of a week may be unreasonable for many sports.[47]

Acute: Chronic Workload Ratio

In 2016, Blanch and Gabbett [17] introduced the Acute: Chronic Workload Ratio (ACWR), which is the most frequently used method of modifying training load before analysing the effect of training load on injury risk.[22 48] The training load on the current week (Day 6 up to Day 0) is considered the "acute" training load. The "chronic" training load is typically defined as the rolling average of the current week and the previous three weeks (Day 27 up to Day 0), known as the or 7:28 ACWR. As shown in,[49] the basic ACWR calculation is:

$$\mathrm{ACWR} = \frac{Acute\ Week}{Chronic\ Weeks * 0.25} = \frac{W_k}{(W_{k-3} + W_{k-2} + W_{k-1} + W_k) * 0.25}$$

Where k is the current week. In the same manner as the rolling average, the traditional ACWR calculation moves iteratively from one week to the next. We calculated ACWR from one day to the next, a calculation less wasteful of data.[47]

ACWR can be calculated in many different ways.[22 23] The time windows for the acute and chronic periods are at the user's discretion.[22 47] The acute load is typically the sum of training load exposures on the current week, but the chronic load can by calculated by either the rolling average or the EWMA.[23 36 50] Finally, in the traditional ACWR, the acute load is included in the denominator. This is known as the "coupled" ACWR. The "uncoupled" ACWR – where the acute load is *not* included in the denominator – has been recommended as a more concrete measure of the change in training load.[18 21] For this simulation study, we chose the coupled 1-week absolute sum: 4 week rolling average ACWR, the most common form of calculation.[23]

The advantage of the ACWR is addressing the potential effect of the relative training load, while also accounting for past exposure. The properties of the ACWR has been explored extensively, with multiple critiques.[18 19 22 23 51] Like EWMA, ACWR needs a completed time window before the first calculation.

Distributed lag non-linear model

The ability of the distributed lag non-linear model (DLNM) to uncover the effect of relative training load was also assessed. The exposure-response function $f(\%\Delta x)$ was assumed to be linear, the same assumption as for the ACWR and week-to-week percentage change. The lag-response function $w(x, l)$ was modelled with restricted cubic splines using 3 knots under all scenarios.

Step 4 Calculating performance measures

Metrics for comparing the model fit, accuracy and certainty of the models were calculated in the final step.

Root-Mean-Squared Error

For a measure of accuracy, we calculated the difference between the predicted cumulative hazard $\hat{\theta}$ and the true cumulative hazard θ used to simulate the survival data for a range of training load values, the absolute bias. The main performance measure was the Root-Mean-Squared Error (RMSE), calculated by:

$$RMSE = \sqrt{mean\left(\left(\hat{\theta} - \theta\right)^2\right)} = \sqrt{mean(bias^2)}$$

RMSE is a combined measure of accuracy and precision, where the lower the RMSE, the better the method.[13] The scale of the RMSE depends on the scale of the coefficients in question, and it is therefore only interpretable by comparing values in the same analysis – the values cannot be interpreted in isolation.[52]

For the relative training load, the ACWR and the week-to-week percentage change methods modified the training load values to a different scale than the one used to simulate the data. The RMSE for the predicted vs. true cumulative hazard, a measure of external validation, could therefore not be calculated for each level of percentage change in training load. Therefore, we also calculated RMSE on the predicted injury value vs. the observed value (the model residuals), as an internal validation:

$$RMSE_{Internal} = \sqrt{mean(residuals^2)}$$

Model fit
Model fit was measured by Akaike's Information Criterion (AIC) which has shown to be more appropriate than BIC for comparison of time-lag models.[27] The AIC can be used to compare non-nested models,[53-55] but the AIC is not comparable if models are run on different sample sizes.[53] Since some methods – EWMA, ACWR – required the completion of a full time period before first calculation, the first 27 rows were removed from the dataset for all methods before fitting the Cox regression model to ensure comparability of the AIC.

Coverage
Coverage was calculated as the proportion of 95% confidence intervals that contained the true value. Average width (AW) of the 95% confidence intervals was also calculated, as a measure of statistical efficiency.

Number of simulations

Using formulas listen in Morris, et al. [52], accepting a Monte Carlo Standard Error of no more than 0.5, the number of simulations needed for an accurate determination of coverage was:

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

The number of simulations needed for an accurate estimate of bias was calculated by:

$$n_{sim} = \frac{s^2}{0.5^2}$$

Where s is the sample variance of bias.[52] For an estimation of variance, a pilot of 200 simulations were run for each constructed relationship. The highest variance in bias was

6.63, and the number of simulations needed to achieve the target MCSE was 176. Since coverage required more simulations to achieve target MCSE, simulation steps 1–4 outlined above were repeated 1 900 times. The mean of each performance measure was calculated across the 1 900 simulations.

IMPLEMENTATION IN A HANDBALL COHORT

The distributed lag non-linear model (DLNM) was implemented on an observed handball cohort to illustrate how it can be used in practice. To explore the potential for a time-dependent, cumulative effect of training load on injury risk, we chose the Norwegian elite youth handball data. The 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.[56]

RPE and duration was collected from the players after each training and match, from which daily sRPE was determined.[56] Timeliness was relatively poor; 53% of activity prompts were answered on the same day, and the mean number of days from prompt to reply was 0.7 (SD = 1.6). Of 47 651 activity prompts, 64% were missing, likely under the missing at random or missing not at random mechanism.[57] Missing sRPE data had previously been imputed with multiple imputation using predicted mean matching,[5] before the data were anonymized.[14] 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,[58] but was not itself imputed before analysis.[59] The duration and RPE variables, the factors from which sRPE is derived, were not included in the imputation model for predicting sRPE.[5] The number of imputed datasets, five, is recommended in most cases.[60] The observed distribution was maintained in the imputed values; therefore the imputation was deemed valid.[14] Although the poor data quality rendered the handball data unsuitable for a study of causal inference, it had a sufficient number of injuries for the current methodology study (n = 472), and previously showed a potential non-linear relationship between training load and injury risk.[14]

The handball players reported whether they had “no health problem”, “a new health problem”, or “an exacerbation of an existing health problem” each day. Any response of “a new health problem” was considered an injury event in the current study. Players were encouraged to report all physical complaints, irrespective of their consequences on sports participation or the need to seek medical attention.[61]

A Cox regression model was run with injury (yes/no) as the outcome and the DLNM of sRPE as the exposure of interest.[62] DLNM combines a dose-function on the magnitude of sRPE, and a lag-function on the distance since Day 0, up to lag 27 (4 weeks). The dose-function was modelled with a restricted cubic splines with 3 knots.[14] Based on AIC, a linear model was chosen for the lag-function. The Cox model was adjusted for sex and age as potential confounders. A frailty term with a gamma distribution was used to account for recurrent events.[12] The model predictions were visualized to assess the ability of DLNM to explore effects. Predictions from each of the imputed datasets were averaged, then visualized.[63]

DATA TOOLS

The simulations were run on an Intel(R) Core(TM) i7-6700K 4.00GHz CPU, 16 GB RAM computer. All statistical analyses and simulations were performed using R 4.1.2[64] with RStudio version 1.4.1717. A GitHub repository is available with R code and data used in the simulations.[65] PermAlgo was used to simulate survival data.[42 66] The slider package was used for calculations on sliding windows,[67] using zoo[68] for rolling averages and TTR[69] for EWMA. Handling time-lag data and performing distributed lag non-linear models was done with DLNM.[70]

ETHICS

Data collection for both 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 Premier League football (722773); Norwegian elite youth handball (407930). All participants provided informed written consent. They were all above the age of 15 and parental consent was not required. Ethical principles were followed in accordance with the Declaration of Helsinki,[71] with the exception that the study was not registered in a publicly accessible database before recruitment of the first subject (a violation of principle number 35). Data were anonymised according to guidelines outlined by The Norwegian Data Protection Authority.[72] The datasets cannot be joined.

REFERENCES

1. O'Kelly M, Anisimov V, Campbell C, et al. Proposed best practice for projects that involve modelling and simulation. *Pharmaceutical statistics* 2017;16(2):107-13.
2. Shmueli G. To explain or to predict? *Statistical science* 2010;25(3):289-310.
3. 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.
4. Rønneberg KL. Seasonal training load quantification in men's Norwegian premier league football: Differences in measured external-and internal training load within microcycles and throughout the competition phase. Norwegian School of Sport Sciences, 2020.
5. Bache-Mathiesen LK, Andersen TE, Clarsen B, et al. Handling and reporting missing data in training load and injury risk research. *Science and Medicine in Football* 2021:1-13. doi: 10.1080/24733938.2021.1998587
6. Curran-Everett D, Williams CL. Explorations in statistics: the analysis of change. *Advances in physiology education* 2015;39(2):49-54.
7. 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.
8. Ekstrand J, Krutsch W, Spreco A, et al. Time before return to play for the most common injuries in professional football: a 16-year follow-up of the UEFA elite Club injury study. *British journal of sports medicine* 2020;54(7):421-26.
9. Bahr R. Why screening tests to predict injury do not work—and probably never will…: a critical review. *Br J Sports Med* 2016;50(13):776-80.
10. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II-binary and time-to-event outcomes. *Statistics in medicine* 2019;38(7):1276-96.
11. Meeuwisse WH, Tyreman H, Hagel B, et al. A dynamic model of etiology in sport injury: the recursive nature of risk and causation. *Clinical Journal of Sport Medicine* 2007;17(3):215-19.
12. Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014;48(17):1287-93.

13. Gasparrini A. Modelling lagged associations in environmental time series data: a simulation study. *Epidemiology (Cambridge, Mass)* 2016;27(6):835.
14. 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(3):e001119. doi: 10.1136/bmjsem-2021-001119
15. Gabbett TJ. The training—injury prevention paradox: should athletes be training smarter and harder? *British journal of sports medicine* 2016;50(5):273-80.
16. 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.
17. 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.
18. 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.
19. Impellizzeri FM, Tenan MS, Kempton T, et al. Acute: Chronic Workload Ratio: Conceptual Issues and Fundamental Pitfalls. *International Journal of Sports Physiology and Performance* 2020;15(6):907-13.
20. Wang C, Stokes T, Steele R, et al. 118 Application of the acute:chronic workload ratio in children. *British Journal of Sports Medicine* 2021;55(Suppl 1):A47-A48. doi: 10.1136/bjsports-2021-IOC.110
21. Wang C, Stokes T, Vargas JT, et al. Injury risk increases minimally over a large range of the acute-chronic workload ratio in children. *American Journal of Epidemiology* 2021 doi: 10.1093/aje/kwab280
22. 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.
23. Wang A, Healy J, Hyett N, et al. A systematic review on methodological variation in acute: chronic workload research in elite male football players. *Science and Medicine in Football* 2021;5(1):18-34.
24. Udby CL, Impellizzeri FM, Lind M, et al. How has workload been defined and how many workload-related exposures to injury are included in published sports injury articles? A scoping review. *Journal of orthopaedic & sports physical therapy* 2020;50(10):538-48.
25. Moussa I, Leroy A, Sauliere G, et al. Robust Exponential Decreasing Index (REDI): adaptive and robust method for computing cumulated workload. *BMJ Open Sport & Exercise Medicine* 2019;5(1)
26. Bhaskaran K, Gasparrini A, Hajat S, et al. Time series regression studies in environmental epidemiology. *International journal of epidemiology* 2013;42(4):1187-95.
27. Gasparrini A. Modeling exposure–lag–response associations with distributed lag non-linear models. *Statistics in medicine* 2014;33(5):881-99.
28. Menaspà P. Are rolling averages a good way to assess training load for injury prevention? *Br J Sports Med* 2017;51(7):618-19.
29. Drew MK, Blanch P, Purdam C, et al. Yes, rolling averages are a good way to assess training load for injury prevention. Is there a better way? Probably, but we have not seen the evidence. *British Journal of Sports Medicine* 2017;51(7):618-19. doi: 10.1136/bjsports-2016-096609
30. 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 Medicine* 2020:1-23.
31. Nakaoka G, Barboza SD, Verhagen E, et al. The association between the acute: chronic workload ratio and running-related injuries in Dutch runners: a prospective cohort study. *Sports medicine* 2021:1-11.

32. Arazi H, Asadi A, Khalkhali F, et al. Association Between the Acute to Chronic Workload Ratio and Injury Occurrence in Young Male Team Soccer Players: A Preliminary Study. *Frontiers in Physiology* 2020;11
33. Xiao M, Nguyen JN, Hwang CE, et al. Increased Lower Extremity Injury Risk Associated With Player Load and Distance in Collegiate Women's Soccer. *Orthopaedic Journal of Sports Medicine* 2021;9(10):23259671211048248.
34. Enright K, Green M, Hay G, et al. Workload and Injury in Professional Soccer Players: Role of Injury Tissue Type and Injury Severity. *Int J Sports Med* 2020;41(02):89-97. [published Online First: 04.12.2019]
35. Schwartz J, Spix C, Touloumi G, et al. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. *Journal of Epidemiology & Community Health* 1996;50(Suppl 1):S3-11.
36. Murray NB, Gabbett TJ, Townshend AD, et al. Calculating acute: chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. *Br J Sports Med* 2017;51(9):749-54.
37. Sedeaud A, De Larochelambert Q, Moussa I, et al. Does an Optimal Relationship Between Injury Risk and Workload Represented by the "Sweet Spot" Really Exist? An Example From Elite French Soccer Players and Pentathletes. *Frontiers in Physiology* 2020;11:1034.
38. Almon S. The distributed lag between capital appropriations and expenditures. *Econometrica: Journal of the Econometric Society* 1965:178-96.
39. Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology* 2006:624-31.
40. Elliott P, Shaddick G, Wakefield JC, et al. Long-term associations of outdoor air pollution with mortality in Great Britain. *Thorax* 2007;62(12):1088-94.
41. Abrahamowicz M, Beauchamp ME, Sylvestre MP. Comparison of alternative models for linking drug exposure with adverse effects. *Statistics in medicine* 2012;31(11-12):1014-30.
42. Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. *Statistics in medicine* 2009;28(27):3437-53.
43. Rogalski B, Dawson B, Heasman J, et al. Training and game loads and injury risk in elite Australian footballers. *Journal of science and medicine in sport* 2013;16(6):499-503.
44. Cousins BE, Morris JG, Sunderland C, et al. Match and Training Load Exposure and Time-Loss Incidence in Elite Rugby Union Players. *Frontiers in Physiology* 2019;10:1413.
45. Ryan MR, Napier C, Greenwood D, et al. Comparison of different measures to monitor week-to-week changes in training load in high school runners. *International Journal of Sports Science & Coaching* 2021;16(2):370-79.
46. Ramskov D, Rasmussen S, Sørensen H, et al. Interactions between running volume and running pace on injury occurrence in recreational runners: A secondary analysis. *Journal of Athletic Training* 2021 doi: 10.4085/1062-6050-0165.21
47. 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.
48. Eckard TG, Padua DA, Hearn DW, et al. The relationship between training load and injury in athletes: a systematic review. *Sports medicine* 2018;48(8):1929-61.
49. Lolli L, Batterham AM, Hawkins R, et al. Mathematical coupling causes spurious correlation within the conventional acute-to-chronic workload ratio calculations. *British Journal of Sports Medicine* 2019;53(15):921-22. doi: 10.1136/bjsports-2017-098110
50. Griffin A, Kenny IC, Comyns TM, et al. The Association Between the Acute: Chronic Workload Ratio and Injury and its Application in Team Sports: A Systematic Review. *Sports Medicine* 2019:1-20.
51. Zouhal H, Boullosa D, Ramirez-Campillo R, et al. Acute: Chronic Workload Ratio: Is There Scientific Evidence? *Frontiers in Physiology* 2021;12

52. Morris TP, White IR, Crowther MJ. Using simulation studies to evaluate statistical methods. *Statistics in medicine* 2019;38(11):2074-102.
53. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic biology* 2004;53(5):793-808.
54. Anderson D, Burnham K. Aic myths and misunderstandings 2006 [Available from: https://sites.warnercnr.colostate.edu/anderson/wp-content/uploads/sites/26/2016/11/AIC-Myths-and-Misunderstandings.pdf accessed 2022-02-22.
55. Glatting G, Kletting P, Reske SN, et al. Choosing the optimal fit function: comparison of the Akaike information criterion and the F-test. *Medical physics* 2007;34(11):4285-92.
56. 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
57. Borg DN, Nguyen R, Tierney NJ. Missing Data: Current Practice in Football Research and Recommendations for Improvement. *Science and Medicine in Football* 2021:1-6. doi: 10.1080/24733938.2021.1922739
58. 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.
59. 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.
60. Van Buuren S. Flexible imputation of missing data: Chapman and Hall/CRC 2018.
61. 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.
62. 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.
63. Miles A. Obtaining predictions from models fit to multiply imputed data. *Sociological Methods & Research* 2016;45(1):175-85.
64. Team RC. R: A language and environment for statistical computing.: Foundation for Statistical Computing, Vienna, Austria, 2021.
65. Bache-Mathiesen LK. Modelling the cumulative effect of training load on injury risk: R code repository 2022 [1.0.0:[Available from: https://github.com/lenakba/modelling-training-load.
66. Sylvestre MP, Evans T, MacKenzie T, et al. PermAlgo: Permutational algorith to generate event times conditional on a covariate matrix including time-dependent covariates. *R package version 11* 2015
67. slider: Sliding Window Functions 0.2.1 [program]: The Comprehensive R Archive Network, 2021.
68. Zeileis A, Grothendieck G. zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software* 2005;14(6):1–27. doi: 10.18637/jss.v014.i06
69. Ulrich J. TTR: Technical Trading Rules. *R package version 0242* 2020
70. Gasparrini A. Distributed lag linear and non-linear models in R: the package dlnm. *Journal of statistical software* 2011;43(8):1.
71. Association GAotWM. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *The Journal of the American College of Dentists* 2014;81(3):14-18.
72. 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 2021-10-14.