1-1-2019

Time-to-event analysis for sports injury research part 1: Time-varying exposures

Rasmus Oestergaard Nielsen

Michael Lejbach Bertelsen

Daniel Ramskov

Merete Møller

Adam Hulme

See next page for additional authors

10.1136/bjsports-2018-099408

Nielsen, R. O., Bertelsen, M. L., Ramskov, D., Møller, M., Hulme, A., Theisen, D., ... & Parner, E. T. (2019). Time-to-event analysis for sports injury research part 1: time-varying exposures. *British Journal of Sports Medicine*, *53*(1), 61-68.
Available here.

Authors

Rasmus Oestergaard Nielsen, Michael Lejbach Bertelsen, Daniel Ramskov, Merete Møller, Adam Hulme, Daniel Theisen, Caroline F. Finch, Lauren Victoria Fortington, Mohammad Ali Mansournia, and Erik Thorlund Parner

Time-to-event analysis for sports injury research part 1: time-varying exposures

Rasmus Oestergaard Nielsen,[1] Michael Lejbach Bertelsen,[1] Daniel Ramskov,[1,2] Merete Møller,[3] Adam Hulme,[4] Daniel Theisen,[5] Caroline F Finch,[6] Lauren Victoria Fortington,[6,7] Mohammad Ali Mansournia,[8,9] Erik Thorlund Parner[10]

ABSTRACT

Background *'How much change in training load is too much before injury is sustained, among different athletes?'* is a key question in sports medicine and sports science. To address this question the investigator/practitioner must analyse exposure variables that change over time, such as change in training load. Very few studies have included time-varying exposures (eg, training load) and time-varying effect-measure modifiers (eg, previous injury, biomechanics, sleep/stress) when studying sports injury aetiology.

Aim To discuss advanced statistical methods suitable for the complex analysis of time-varying exposures such as changes in training load and injury-related outcomes.

Content Time-varying exposures and time-varying effect-measure modifiers can be used in time-to-event models to investigate sport injury aetiology. We address four key-questions (i) Does time-to-event modelling allow change in training load to be included as a time-varying exposure for sport injury development? (ii) Why is time-to-event analysis superior to other analytical concepts when analysing training-load related data that changes status over time? (iii) How can researchers include change in training load in a time-to-event analysis? and, (iv) Are researchers able to include other time-varying variables into time-to-event analyses? We emphasise that cleaning datasets, setting up the data, performing analyses with time-varying variables and interpreting the results is time-consuming, and requires dedication. It may need you to ask for assistance from methodological peers as the analytical approaches presented this paper require specialist knowledge and well-honed statistical skills.

Conclusion To increase knowledge about the association between changes in training load and injury, we encourage sports injury researchers to collaborate with statisticians and/or methodological epidemiologists to carefully consider applying time-to-event models to prospective sports injury data. This will ensure appropriate interpretation of time-to-event data.

[1]Section for Sports Science, Department of Public Health, Aarhus University, Aarhus, Denmark
[2]Department of Physiotherapy, University College Northern Denmark, Aalborg, Denmark
[3]Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
[4]Centre for Human Factors and Sociotechnical Systems, Faculty of Arts, University of the Sunshine Coast, Maroochydore DC, Queensland, Australia
[5]Sports Medicine Research Laboratory, Luxembourg Institute of Health, Luxembourg, Luxembourg
[6]Australian Centre for Research into Injury in Sport and its Prevention, School of Medical and Health Sciences, Edith Cowan University, Perth, Western Australia, Australia
[7]Faculty of Science and Technology, Federation University Australia, Ballarat, Victoria, Australia
[8]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran
[9]Sports Medicine Research Center, Neuroscience Institute, Tehran University of Medical Sciences, Tehran, Iran
[10]Section for Biostatistics, Department of Public Health, Aarhus University, Aarhus, Denmark

Correspondence to Dr Rasmus Oestergaard Nielsen, Section for Sports Science, Department of Public Health, Aarhus University, Aarhus 8000, Denmark; roen@ph.au.dk

INTRODUCTION

In the past decades, general methodologists of science insisted that it was impossible to measure how health-related exposures and outcomes changed over time.[1] Rather, researchers interested in the study of change were encouraged to 'frame their questions in other ways'.[1] Clearly this was a poor advice. Today, one of the overarching goals of sports injury research is to understand why, and when, athletes sustain injury. A current example of this is the 'too much, too soon' theory.[2]

This theory proposes that athletes are at greater risk of injury following a sudden change in training load, a sudden change in the magnitude of the load, a sudden change in the way the load is distributed, or a combination of these changes.[3] Clearly, the study of change over time is crucial in the sport injury context. Refining the concept, the sports injury community has become increasingly interested in research questions such as '*how much change in training load is too much before injury is sustained, among athletes with different characteristics?*'[3–5] as it is now feasible to measure how individual training loads change over time, using wearable devices with Global Positioning System to facilitate cost-efficient data collection. Across many sports, researchers collect longitudinal data, for example, on training load and injury occurrence, from hundreds of athletes over a full season, a full year, or ideally, longer periods.[5] So, how should these data be analysed to study the impact of changes over time on the development of sports injury?

To study change over time, time-varying *exposures* (eg, change in training load) and time-varying *outcomes* (eg, change in injury status) are two essential concepts.[6] With these concepts in mind, analysing the association between changes in training load and the onset of injury has received careful attention in the scientific literature. In fact, a plethora of original articles[5] and consensus reports[2 7] were published during the years 2015–2017 alone. In a systematic review[5] of 31 articles that examined the training load-injury relationship, X^2tests and logistic regression[8 9] were identified as the most commonly used analytical approaches, whereas other approaches such as time-to-event models were used in only two articles (6%)[10 11] (table 1). These findings have been confirmed in a recent methodological paper as less than 10% of all results in the identified studies were based on time-to-event or multilevel modelling.[12]

The importance of this finding is that, unless the time to injury is discretised (eg, transferring continuous timescales or variables into units for analyses), it is *not* possible to include time-varying exposures in traditional logistic regression models or X^2 tests. Since logistic regression has been the primary choice of many sports injury researchers, this initiates an important debate: How well has the 'too much change in training load, too soon' theory been explored in the existing literature? And how reliable are the results that have been reported to date? To facilitate more refined insights into sports injury occurrence,[2] we should carefully consider which analytical approach best assesses associations between changes in training load and injury onset.

Digging deeper into the concept of time-varying exposures, sports injury researchers (particularly methodologists, biostatisticians and epidemiologists) can learn from the broad biostatistical subdiscipline called *time-to-event modelling*. Time-to-event modelling allows analysis of changes in training load and

Table 1 Overview of the statistical methods, as stated by the authors in the statistics section in each manuscript, used to examine the association between training load and injury in 31 original articles included in a systematic review by Drew and Finch[5]

First author (reference number)	Year	Sample size	Statistical method*	Supplementary methods
Rugby and rugby union				
Gabbett[49]	2004	79	X^2	One-way analysis of variance
Gabbett[50]	2004	220	X^2	One-way analysis of variance
Gabbett[51]	2005	153	GLM (OR)	
Gabbett[52]	2007	183	Logistic regression	X^2
Brooks[53]	2008	502	Pearson correlation	
Gabbett[54]	2010	91	Logistic regression	Two-way analysis of variance
Killen[55]	2010	36	X^2	
Gabbett[56]	2011	79	Pearson correlation	
Gabbett[11]	**2012**	**34**	**Time to event (Cox)**	
Hulin[57]	2016	53	Logistic regression	
Cross[58]	2016	173	GLM	
Cricket				
Dennis[59]	2003	90	2×2 tables (risk ratio)	T-test
Dennis[60]	2005	44	2×2 tables (risk ratio)	
Orchard[61]	2009	129	2×2 tables (risk ratio)	T-test
Hulin[62]	2014	28	Logistic regression	
Orchard[63]	2015	235	Logistic regression	
Football				
Lovell[64]	2006	19	Logistic regression	
Piggot (master's thesis)	2009	16	Pearson correlation	
Brink[65]	2010	53	Multinomial regression	
Rogalski[66]	2013	46	Logistic regression	X^2
Colby[67]	2014	46	Logistic regression	X^2
Ehrmann	2015	19	Unable to assess article	
Other sports				
Lyman[68]	2001	398	GLM	
Lyman[69]	2002	476	GLM	Logistic regression
Anderson[70]	2003	12	Pearson correlation	
Wilson[71]	2010	20	X^2	Pearson correlation
Visnes[72]	2013	141	Logistic regression	
Wheeler[73]	2013	7	Residual maximum likelihood	
Malisoux[10]	**2013**	**154**	**Time to event (Cox)**	
Bahr[74]	2014	44	Unknown model	
Hellard[75]	2015	28	Logistic regression	

Papers in which time-to-event models were used are highlighted in bold.
Cox refers to Cox proportional hazards regression.
Football includes soccer and Australian football. Other sports: rowing, baseball, basketball, swimming, volleyball and multiple sports.
*Used to examine the association between training load and sports injury.
GLM, generalised linear model.

their relationship to sports injury.[6] These analyses bring novel insights for sports researchers, coaches, athletes and clinicians, but come at the price of requiring more advanced statistical skills.[5 13 14] Educational articles, targeting sports injury researchers, to explain the potential application of time-to-event analysis are required to facilitate their uptake. This process of translating complex statistical models and methodological concepts to applied users has already begun with a series of articles having been published in *BJSM* and related journals.[6 15 16] In our 2016 paper in *Journal of Orthopaedic and Sports Physical Therapy* (*JOSPT*),[6] we focused on different measures of association, such as cumulative relative risk, cumulative risk difference and the classical hazard rate ratio, and we shared tips on how to interpret the statistical results. We introduced the more advanced concepts of time-varying exposures and time-varying outcomes.[6] However, we did not detail the importance of time-varying exposures in relation to changes in training load. In addition, there have been important advances in time-to-event models recently reported in technical statistics papers.[17 18] Consequently, here we aim to provide accessible, non-mathematical descriptions to help the interested *BJSM* community member better understand time-varying exposures. In a non-technical language, we present a range of statistical methods and tools for the analysis of exposure variables that change over time, such as change in training load, which have been developed by experts.[19]

The purpose of this paper is to discuss how changes in training load, which is an example of a time-varying exposure, can be used in time-to-event models to investigate injury aetiology in the sport context. This is part of a strategic editorial commitment by *BJSM*[4 19–23] and other journals[6 24 25] to advance the quality of methods used in sports medicine research. In this article, part 1 of two articles, we focus on time-varying exposures. Table 2 provides a brief overview of the key questions addressed in this manuscript. *Time-varying outcomes*, competing risks and subsequent injuries, through a time-to-event lens, are presented in an accompanying paper entitled '*Time-to-event Analysis for Sports Injury Research Part 2: Time-varying Outcomes*'.

TIME-TO-EVENT AND TIME-VARYING EXPOSURES

One critical feature of prospective, longitudinal sports injury data is that exposures, for example, training patterns, strength, flexibility and behaviour (risk-taking), inevitably change between the time they were measured at baseline and during the follow-up period.[6] Recognising and incorporating such changes into an analysis is required if sports injury aetiology is to be examined. Still, the following questions remain: (1) Does time-to-event modelling allow researchers to include change in training load as a time-varying exposure to sport injury development? (see: *Time-varying exposures question 1*); (2) Why is time-to-event analysis superior to other analytical concepts? (see: *Time-varying exposures*

Table 2 Four common questions in sports injury research and how they can be addressed using time-to-event analysis

Time-to-event analysis and time-varying exposures: four common questions in sports injury research	
Question 1 *What is a time-varying exposure?*	**Sports injury researchers are addressing the question '*How much change in training load is too much before injury is sustained, among athletes with different characteristics?*' Does time-to-event modelling allow them to include change in training load as a time-varying exposure to sport injury development?** **Key point 1:** Time-to-event modelling is well suited to deal with time-varying exposures and its association with sports injury. When using a time-varying training load exposure, the primary exposure of interest must be labelled '*change* in training load', not 'training load'. Sudden spikes and reductions in training load are not exposure variables, but exposure levels (known as 'states'). Consequently, researchers do not examine the association between sudden spikes in training load and injury. They examine the association between changes in training load and injury.
Question 2 *Why time-to-event modelling?*	**Why is time-to-event analysis superior to other analytical concepts when analysing training load-related data that changes status over time?** **Key point 2:** Unlike logistic regression and X^2 test, time to event allows flexibility for the sports injury researcher to take into account censoring and compare injury risk across time-varying exposures by using delayed entry functions.
Question 3 *Considerations regarding training load-related time-varying exposures*	**How should sports injury researchers include change in training load in a time-to-event analysis?** **Key point 3a:** In a time-to-event analysis, change in training load can be included as a categorised variable. This enables for examining non-linear dose–response relationships in the association between changes in training load and sports injury. **Key point 3b:** The acute:chronic workload ratio (ACWR) concept is a concrete example of a categorised time-varying training load. In this, the cut-off points of 0.8 and 1.3 have been suggested. It must be stressed that the most suitable value(s) for the cut-off points and whether changes in training load should be based on absolute and/or relative changes remain areas of uncertainty and discussion. **Key point 3c:** The sports injury researcher can examine changes in training load using either fixed states and/or transitions between states. Therefore, sports injury researchers are advised to specify up front whether the main objective of the study is to examine injury risk in relation to (1) different (but constant) workload states (eg, low, medium/sweet, high), or (2) the transitions between workload states (eg, moving from high to medium).
Question 4 *What about other time-varying exposures?*	**Many other factors than training load also change status over time (eg, body mass, strength, flexibility). Are sports injury researchers able to include such variables into the time-to-event analysis?** **Key point 4:** Yes. Many variables change over a player's season or career that can be important to consider in respect to sports injury. These variables can be included as time-varying effect-measure modifiers and/or time-varying confounders. The challenge in this scenario is to have sufficient data to support inclusion of these variables.

question 2); (3) How can the association between changes in training load and injury be examined using time-to-event models? (see: *Time-varying exposures question 3*); and (4) How many different types of time-varying exposures can be included in time-to-event modelling (see: *Time-varying exposures question 4*).

Time-varying exposures question 1: What are time-varying exposures?

Imagine that you wish to address the question '*how much change in training load is too much before injury is sustained, among athletes with different characteristics?*' in your upcoming research project. How does time-to-event modelling allow you to include change in training load as a time-varying exposure to sport injury development?

The concept of change is important as training patterns and athletic participation fluctuate on a monthly, weekly or even a daily basis. As changes in training load vary over time, this variable is a so-called *time-varying exposure*. Time-to-event modelling is well suited to deal with time-varying exposures and its association with sports injury.

In sports science, we assume training spikes (eg, excessive progression in training load) can lead to sports injury. Still, we also acknowledge that athletes have recovery periods (reduced training load). Consequently, the researcher needs to include a time-varying exposure variable that consists of sudden spikes and slight increases and/or reductions in training load over time. Therefore, it is more appropriate to label the time-varying exposure of interest as '*change in training load*' rather than 'sudden spikes' or 'workload progression'. The latter two are levels (so-called 'exposure states') of exposure, not the exposure variable, which is change in training load. Accordingly, the researcher examines the association between changes in training load and sports injury.

If the researcher leaves out 'change in' and specifies that she/he examines the association between training load and injury, the time-varying nature of the exposure is not clearly specified. In the literature, there are several examples of studies using time-fixed training load-related exposures. For instance, Walter *et al*[26] examined the association between weekly mileage and running injury. The women running more than an *average* mileage per week over a 3-month period exceeding 40 had an increased injury risk of 242% compared with those running below 10 miles/week on average in the preceding 3 months. This enabled Walter *et al* to identify a subgroup of female runners at increased injury risk. However, it remains open to speculation *why* the women exceeding an average of 40 miles/week over a 3-month period were more vulnerable to injury. If injury occurs owing to sudden changes in one or more variables, we need to consider these sudden changes.[3] Here, the concept of time-varying exposures is a necessary ingredient.

Key point 1: Time-to-event modelling is well suited to deal with time-varying exposures and its association with sports injury. When using a time-varying training load exposure, the primary exposure of interest must be labelled 'change in training load', not 'training load'. Sudden spikes and reductions in training load are not exposure variables, but exposure levels (known as 'states'). Consequently, researchers do not examine the association between sudden spikes in training load and injury. They examine the association between changes in training load and sports injury.

Time-varying exposures question 2: why time-to-event modelling?

Regardless of whether changes in training load are defined based on relative changes, for example, the acute:chronic workload ratio (ACWR),[27] a modified version of the ACWR,[10 28 29] biweekly changes,[30] or absolute changes, a high-quality statistical analysis is needed in order to take into account that each athlete can change their status (in statistical terms: transit between states) many times during the course of follow-up. In the present paper, we guide researchers towards time to event when choosing a statistical model. Still, one may speculate: Why is time-to-event analysis superior to other analytical concepts when analysing training load-related data that changes status over time?

In prospective studies, researchers collect data from the same individuals over time. Most likely, sports injury researchers will experience that the training load data from each individual will vary from the time of inclusion (baseline) and during the follow-up. Having such data allows sports injury researchers to model and *compare* the injury risks across athletes being in different training 'zones' or 'levels' (statistically: states) over time. For instance, one could examine if injury risk is greater following a sudden increase in training load compared with a slight increase in training load. Here, simple logistic regression and X^2 test are too restricted to be able to provide answers. Advanced analytical techniques are required. In time-to-event models, the researcher is able to compare injury risk across different changes in training load using what is termed a 'delayed entry function'. Then, an individual should only be considered at risk in the time period at which the individual is in the given state. Knowing the concept of delayed entry is important as it allows the sports injury researcher to deliver a specific request to the statistician: "*We need to analyse changes in each of the athletes'/players' training load data using delayed entry. Can you help me with that please?*"

In addition to acknowledging the importance of dealing with the delayed entry, to model how the injury risk depends on change in training load, time-to-event models allow the possibility of censoring participants as it is likely in prospective studies that some study participants leave the study during follow-up for various reasons. If the researchers do not take into account censoring, they assume that all participants in the study complete follow-up. This is also a very speculative assumption.

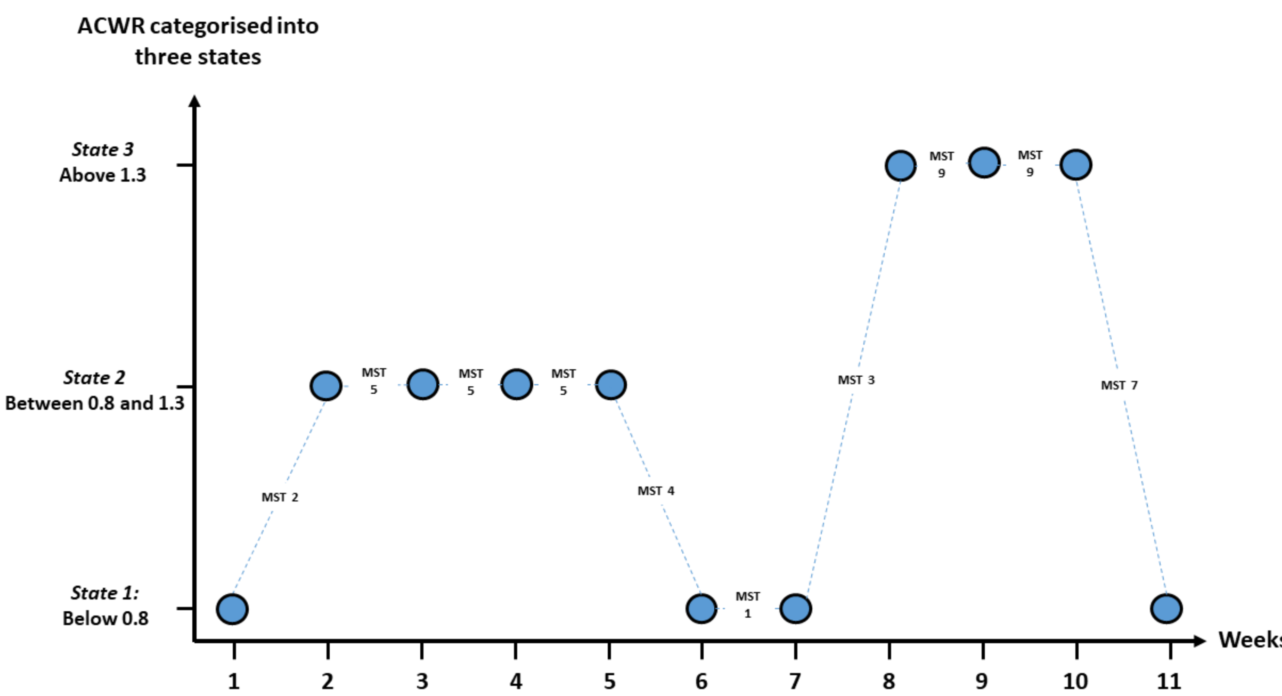

Figure 1 Overview of the concepts of acute:chronic workload ratio (ACWR) states and multistate transitions (MST) using an n=1 athlete example. Imagine that we register the ACWR of one athlete during an 11-week follow-up. On the y-axis, the ACWR was classified into one of the three following states *each* week during the 11-week follow-up (marked with blue circle): state 1: below 0.8; state 2: between 0.8 and 1.3; and state 3: above 1.3. Then, the athlete is able to move/switch/transit between these states *between* each week. Consequently, the following nine MSTs are possible in the example: MST1: below 0.8 and remaining below 0.8; MST2: below 0.8 to between 0.8 and 1.3; MST3: below 0.8 to above 1.3; MST4: between 0.8 and 1.3 to below 0.8; MST5: between 0.8 and 1.3 and remaining between 0.8 and 1.3; MST6: between 0.8 and 1.3 to above 1.3; MST7: above 1.3 to below 0.8; MST8: above 1.3 to between 0.8 and 1.3; MST9: above 1.3 and remaining above 1.3. The concept of states and transitions illustrated is directly transferable to time-varying outcomes (eg, changes in injury severity) and time-varying effect-measure modifiers.

Taking into account varying training data and censoring is possible in time-to-event modelling, and it is possible to include time-varying exposures. These strengths make time-to-event modelling vastly more sophisticated than, for example, logistic regression analysis in which the analysis is limited to the inclusion of time-fixed variables and is unable to use delayed entry. In addition, the censoring of participants in logistic regression requires either adjustment for length of follow-up, which leads to biased estimates, or additional statistical programming because the codes to run the analyses are not included in most readily available software. Consequently, time-to-event modelling should be considered as a preferred analytical strategy in sports injury research when examining the association between changes in training load and injury occurrence—at least when the outcome is dichotomised or categorised.

Key point 2: In contrast to logistic regression and X^2 test, time-to-event modelling allows the researcher to take into account censoring and compare injury risk across time-varying exposures by using delayed entry functions.

It is important to recognise that alternative methods to handle time-varying exposures exist. Other modelling strategies (eg, generalised estimating equations, random effects models or multilevel regression approaches[31]) can be used as they, like time to event, take into account the repeated measures by clustering observations over time within individuals. Consequently, generalised estimating equations, random effects models and/or multilevel regression approaches also provide researchers with opportunities to analyse data based on repeated measurements and with within-subject correlation. In those analyses, the outcome can be categorical or quantitative/continuous. A description of these methods can be found elsewhere, for example, in Smith and Walls[31] study. To the best of our knowledge, readily available software does not allow researchers to deal as easy with the concept of exposure variables, like changes in training load, which changes status over time when using generalised estimating equations, random effects models and/or multilevel regression approaches. Therefore, time-to-event models are a more feasible approach for most sports injury researchers.

Time-varying exposures question 3: how to include changes in training load in a time-to-event analysis?

When the researcher has chosen to include changes in training load as a time-varying exposure variable and use time to event as analytical approach, there several practical options remain:

- *Continuous versus categorised exposure:* A priori one could assume that there is a linear relationship between changes in training load and injury risk. However, it is reasonable to question whether or not a linear dose–response relationship exists in the association between changes in training load and injury.[27] If the relationship is assumed to be non-linear, the next step becomes to categorise time-varying exposure based on a set of a priori defined cut-offs (perhaps three or four 'states' of change). Although spline regression and fractional polynomials have shown promising results when handling continuous training load exposures.[32] Here, we limit ourselves to categories with cut-offs, which allows for the examination of exponentialised relationships and/or U-shaped patterns to describe associations between changes in training load and injury.[27]
  Key point 3a: In a time-to-event analysis, change in training load can be included as a categorised variable. This enables examining non-linear dose–response relationships in the association between changes in training load and sports injury.
- *Defining cut-offs:* Choosing cut-offs will categorise the exposure variable, in this case changes in training load, into certain exposure *states*. Note: the term 'exposure states' and not 'exposure groups'. In exposure groups, each individual is assigned to a certain time-fixed exposure group without the possibility to transit between groups. Importantly, *'state'* indicates that each individual has the opportunity to switch/transit between exposure states during follow-up.
  To define cut-offs, we encourage researchers to use appropriate, up-to-date scientific/biological rationale as the basis for their choices instead of allowing the statistical software program to choose the cut-offs based on data-driven knots. As these data-driven knots are produced by a software program, the approach leaves little room for hypothesis-driven research. Taking the concept of the ACWR as a concrete example of time-varying training load, we know that cut-offs of 0.8 and 1.3 have been suggested as being relevant[27] and in other articles, authors have suggested using load progressions of 10%, 20%, 30% or 60% as indicating critical change in load.[28 30] Although these different ways of categorising changes in training load are appealing to use in future studies, it remains uncertain which cut-offs are appropriate. This leaves the researcher with many possibilities for choosing the cut-offs they believe, based on subject matter knowledge and studies from the literature, are most appealing.
  Key point 3b: The ACWR concept is a concrete example of time-varying training load. Here, the cut-offs of 0.8 and 1.3 have been suggested as important cut-offs. Still, it must be stressed that it is uncertain as to which cut-offs are suitable and if changes in training load should be based on absolute and/or relative changes.
- *Choosing between states and transitions*: Analysts should decide if they want to analyse changes based on exposure states themselves or transitions between exposure states, as these are two different concepts and both are valid. If change in the ACWR is used as the primary exposure of interest, and the corresponding cut-offs are set at 0.8 and 1.3, then the researcher considers three exposure states: below 0.8, between 0.8 and 1.3, and above 1.3 (figure 1). This is appropriate when the researcher aims to examine whether a certain state (eg, the 'sweet spot' between 0.8 and 1.3) is associated with greater/lower injury risk compared with being in another state (below 0.8 or above 1.3).
  On the other hand, asking whether transitions between two states convey changes in injury risk is a question of interest that the previous approach does not answer. Specifically, is it more injurious to switch from the sweet spot (say, 'between 0.8 and 1.3') to an ACWR of >1.3, rather than remaining constantly in the 0.8–1.3 sweet spot. With three exposure states (below 0.8, between 0.8 and 1.3, and above 1.3) there are nine different options for multistate transitions (MST):
  – MST1: Below 0.8 and remaining below 0.8.
  – MST2: Below 0.8 to between 0.8 and 1.3.
  – MST3: Below 0.8 to above 1.3.
  – MST4: Between 0.8 and 1.3 to below 0.8.
  – MST5: Between 0.8 and 1.3 and remaining between 0.8 and 1.3.
  – MST6: Between 0.8 and 1.3 to above 1.3.

- MST7: Above 1.3 to below 0.8.
- MST8: Above 1.3 to between 0.8 and 1.3.
- MST9: Above 1.3 and remaining above 1.3.

Sports injury researchers are advised to specify up front whether the main objective of the study is to examine injury risk in relation to (1) different (but constant) workload states (eg, low, medium/sweet, high), or (2) the transitions between workload states (from one state to another).[33]

Key point 3c: The sports injury researcher can examine changes in training load using either states and/or transitions between states. Therefore, sports injury researchers are advised to specify up front whether the main objective of the study is to examine injury risk in relation to (1) different (but constant) workload states (eg, low, medium/sweet, high), or (2) the transitions between workload states (from one state to another).

INCLUDING ADDITIONAL TIME-VARYING VARIABLES

To this point, we have presented the basics surrounding time-varying training load-related exposures. The following takes the sports injury researcher into the next (advanced) step by considering time-varying effect-measure modification[28] and time-varying confounding.[34]

Time-varying exposures question 4: other time-varying variables

At this stage, the researcher may acknowledge the importance of changes in training load. However, many other factors change status over time (eg, body mass, strength, flexibility). Can one include multiple time-varying exposures/variables into the time-to-event analysis?

In the present article, emphasis is placed on describing training load as a time-varying exposure. However, the occurrence of related sports injuries is highly dynamic in nature,[3 27 35–37] and so it is equally important to understand that other variables may also contribute to injury development. Many of these variables are also time varying, such as equipment usage, body mass, strength, sleep and diet. Consequently, how to handle the way in which other variables change over time has to be considered when using time-to-event modelling approaches. In a handball-related study, Møller *et al* used a time-to-event model to examine the association between changes in training load and shoulder-related injury across subgroups of players with different levels of scapular control.[28] In addition to including changes in training load as a time-varying exposure, scapular control was included as a time-varying effect-measure modifier—not just as a baseline value. This was under the assumption that the neuromuscular function supporting scapular control among handball players will not be constant across one or more seasons—and thus the duration of the study. The biological rationale was that scapular control can change during a season because of either (1) purposeful rehabilitation, (2) muscle inhibition/imbalance over the course of a season because of, for example, muscle fatigue, or (3) frank injury to the rotator cuff (shoulder tendinopathy) or another body part that limits the player's ability to train.

In the handball study above, the data set-up included a time-varying training load exposure and a time-varying non-training-related variable (scapular control). This is an example of the concept of *effect-measure modification*. By using the concept of effect-measure modification, the researcher is able to examine how much change in training load is too much among athletes with different characteristics.

When a researcher is interested in examining the combined effect of two exposures that both are related to change in training load, say change in running distance and change in running intensity, this is also possible via time-to-event modelling using *interaction*, not effect-measure modification or confounding, between the two time-varying training load-related variables. Such an analytical approach is suitable if a researcher wants to examine questions such as: Is it more injurious to progress in running distance *and* in running intensity at the same time compared with progressing in running distance while running at the same intensity? For an example of such a set-up, we guide the reader to look deeper into table 2 in the related article entitled '*Time-to-Event Analysis for Sports Injury Research Part 2: Time-Varying Outcomes*'.

Key point 4: Many risk factors for sports injury, for example, training load, body mass, strength and flexibility, can be included in the same time-to-event analysis as time-varying training load exposures (and the interaction between them), time-varying effect-measure modifiers and/or time-varying confounders.

Other research questions

The goal of many sport injury researchers has been to predict injury risk in subgroups of athletes who present with a certain characteristic (or putative biologic exposure).[4] To do this, researchers need to carefully consider the research question '*How much change in training load is too much before injury is sustained, among athletes with different characteristics?*' and the concepts of time-varying exposures and time-varying effect-measure modification.[38] If the sport injury researcher has the different, but equally important, research question '*What is the average causal effect of body mass on sports injury occurrence?*' then other analytical approaches need consideration, namely: time-varying confounding.

Time-varying confounding

Over the past decade, techniques underpinning causal inference have emerged.[39] Here, the goal is not to investigate stratum-specific differences but to estimate the causal effect of an exposure on sports injury after adjustment for confounding. Confusing the effect of interest with non-causal associations which results from a common cause of both the exposure and outcome.[34] All study designs, including randomised controlled trials, are subject to random confounding and compliance problems.[40 41] A minimally sufficient set of confounders (ie, a set of confounders sufficient for confounding elimination of which no subset is sufficient) must be identified via causal diagrams. Then, one approach could be to adjust for the confounders using standard regression modelling.[42 43] In practice, many exposures of interest are time varying, and the values of potential confounders may change over time leading to time-varying confounding. However, standard regression methods for analysis of longitudinal data such as time-dependent Cox regression do not appropriately adjust for time-varying confounding, and causal methods including inverse probability of treatment weighting, the parametric G-formula and G-estimation, or collectively G-methods should be used instead.[34 44–47]

Recent developments within statistics have opened the use of G-methods in time-to-event analysis using proportion-based measures of association.[17] Here, one needs to address assumptions regarding right censoring, since special techniques to estimate average causal effects are required. The pseudo-observation method has proved valuable for this purpose when applying direct standardisation (G-formula) or inverse probability weights (based on propensity scores).[17] Conclusively, sports injury researchers should be aware that estimation of average causal effects is possible when using Cox regression and a generalised linear model (pseudo-observations). Such analyses are complicated and often require collaboration with a statistician.[48]

Time-varying outcomes

In sports injury research, the concepts of time-varying exposures and outcomes appear to have been rarely used in combination with time-to-event models. Although the reasons for this are unknown, it could be due to either a lack of awareness among sport injury researchers about the potential utility of time-to-event models, or alternatively, the perceived difficulty regarding their use in practice. Another reason could certainly be the limited sample size and event per variable, which are related to the amount of sports injuries available in the data set. This issue will be addressed in the follow-up article entitled '*Time-to-event Analysis for Sports Injury Research Part 2: Time-varying Outcomes*'.

CONCLUSION

Careful attention on how to analyse the time-varying relationship between changes in training load and changes in injury status is needed to address the research question '*How much change in training load is too much, among athletes with different characteristics, before injury is sustained?*' Time-to-event models are suitable for analysing this highly dynamic relationship as they take into account censoring and the within-individual correlation of follow-up data. Naturally, many factors other than training load change status over time. These include, but are not limited to, body mass, equipment usage, sleep and strength. To take into account that player/athlete characteristics change over time, analytical concepts such as time-varying effect-measure modifiers and/or time-varying confounders are important.

Research into sports injuries is undergoing a transformation with increased attention to stronger analytical methods. As these new insights have potential value for sports injury researchers, there is a need to revisit and further elaborate on these analytical concepts. The analytical approaches presented in this paper require specialist knowledge and well-honed statistical skills to master. Cleaning data sets, setting up the data, performing the analyses and interpreting the results are a time-consuming process which requires dedication and, most likely, assistance from methodological peers. To increase knowledge about the association between changes in training load and injury, sports injury researchers are encouraged to collaborate with statisticians and/or methodological epidemiologists to carefully consider applying time-to-event models to their prospective sports injury data and ensure appropriate interpretations of time-to-event data.

Contributor RON drafted the manuscript and all authors revised it for important intellectual content. MAM and EP drafted the section about causal inference.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Obtained.

Ethics approval Local ethics committee in the central Denmark region.

Provenance and peer review Not commissioned; externally peer reviewed.

OPEN ACCESS

To cite Nielsen RO, Bertelsen ML, Ramskov D, *et al*. *Br J Sports Med* 2019;**53**:61–68.

Accepted 13 October 2018
Published Online First 9 November 2018

Linked

► http://dx.doi.org/10.1136/bjsports-2018-100000

Br J Sports Med 2019;**53**:61–68.
doi:10.1136/bjsports-2018-099408

REFERENCES

1 Cronbach LJ, Furby L. How we should measure "change": Or should we? *Psychol Bull* 1970;74:68–80.
2 Soligard T, Schwellnus M, Alonso JM, *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 Bertelsen ML, Hulme A, Petersen J, *et al*. A framework for the etiology of running-related injuries. *Scand J Med Sci Sports* 2017;27:1170–80.
4 Nielsen RO, Bertelsen ML, Møller M, *et al*. Training load and structure-specific load: applications for sport injury causality and data analyses. *Br J Sports Med* 2018;52:1016–7.
5 Drew MK, Finch CF. The relationship between training load and injury, illness and soreness: A systematic and literature review. *Sports Med* 2016;46:861–83.
6 Nielsen RØ, Malisoux L, Møller M, *et al*. Shedding light on the etiology of sports injuries: A look behind the scenes of time-to-event analyses. *J Orthop Sports Phys Ther* 2016;46:300–11.
7 Schwellnus M, Soligard T, Alonso JM, *et al*. How much is too much? (Part 2) international olympic committee consensus statement on load in sport and risk of illness. *Br J Sports Med* 2016;50:1043–52.
8 Malone S, Owen A, Newton M, *et al*. The acute:chonic workload ratio in relation to injury risk in professional soccer. *J Sci Med Sport* 2017;20:561–5.
9 Malone S, Roe M, Doran DA, *et al*. High chronic training loads and exposure to bouts of maximal velocity running reduce injury risk in elite Gaelic football. *J Sci Med Sport* 2017;20:250–4.
10 Malisoux L, Frisch A, Urhausen A, *et al*. Monitoring of sport participation and injury risk in young athletes. *J Sci Med Sport* 2013;16:504–8.
11 Gabbett TJ, Ullah S. Relationship between running loads and soft-tissue injury in elite team sport athletes. *J Strength Cond Res* 2012;26:953–60.
12 Windt J, Ardern CL, Gabbett TJ, *et al*. Getting the most out of intensive longitudinal data: a methodological review of workload-injury studies. *BMJ Open* 2018;8:e022626.
13 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. *Br J Sports Med* 2017;51:618.2–9.
14 Drew MK, Cook J, Finch CF. Sports-related workload and injury risk: simply knowing the risks will not prevent injuries: Narrative review. *Br J Sports Med* 2016;50:1306–8.
15 Bahr R, Holme I. Risk factors for sports injuries-a methodological approach. *Br J Sports Med* 2003;37:384–92.
16 Ullah S, Gabbett TJ, Finch CF. Statistical modelling for recurrent events: an application to sports injuries. *Br J Sports Med* 2014;48:1287–93.
17 Andersen PK, Syriopoulou E, Parner ET. Causal inference in survival analysis using pseudo-observations. *Stat Med* 2017;36:2669–81.
18 Overgaard M, Andersen PK, Parner ET. Regression analysis of censored data using pseudo-observations: An update. *The Stata Journal* 2015;15:809–21.
19 Verhagen E, Stovitz SD, Mansournia MA, *et al*. BJSM educational editorials: methods matter. *Br J Sports Med* 2018;52.
20 Hulme A, Thompson J, Nielsen RO, *et al*. Towards a complex systems approach in sports injury research: simulating running-related injury development with agent-based modelling. *Br J Sports Med*. In Press. 2018. doi: bjsports-2017-098871.
21 Nielsen RO, Bertelsen ML, Verhagen E, *et al*. When is a study result important for athletes. *Br J Sports Med* 2017;51:1454–5.
22 Stovitz SD, Verhagen E, Shrier I. Misinterpretations of the 'p value': a brief primer for academic sports medicine. *Br J Sports Med* 2017;51:1176–7.
23 Stovitz SD, Verhagen E, Shrier I. Distinguishing between causal and non-causal associations: implications for sports medicine clinicians. *Br J Sports Med* 2017. doi: bjsports-2017-098520.
24 Shrier I, Steele RJ, Zhao M, *et al*. A multistate framework for the analysis of subsequent injury in sport (M-FASIS). *Scand J Med Sci Sports* 2016;26:128–39.
25 Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008;8:70–2288.

26 Walter SD, Hart LE, McIntosh JM, *et al*. The Ontario cohort study of running-related injuries. *Arch Intern Med* 1989;149:2561–4.
27 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.
28 Møller M, Nielsen RO, Attermann J, *et al*. Handball load and shoulder injury rate: a 31-week cohort study of 679 elite youth handball players. *Br J Sports Med* 2017;51:231–7.
29 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;51:749–54.
30 Nielsen RØ, Parner ET, Nohr EA, *et al*. Excessive progression in weekly running distance and risk of running-related injuries: an association which varies according to type of injury. *J Orthop Sports Phys Ther* 2014;44:739–47.
31 Smith DM, Walls TA. Multiple time scale models in sport and exercise science. *Meas Phys Educ Exerc Sci* 2016;20:185–99.
32 Carey DL, Crossley KM, Whiteley R, *et al*. Modeling training loads and injuries: The dangers of discretization. *Med Sci Sports Exerc* 2018;50:2267-2276.
33 Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007;26:2389–430.
34 Mansournia MA, Etminan M, Danaei G, *et al*. Handling time varying confounding in observational research. *BMJ* 2017;359:j4587.
35 Meeuwisse WH, Tyreman H, Hagel B, *et al*. A dynamic model of etiology in sport injury: the recursive nature of risk and causation. *Clin J Sport Med* 2007;17:215–9.
36 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.
37 Hulme A, Finch CF. From monocausality to systems thinking: a complementary and alternative conceptual approach for better understanding the development and prevention of sports injury. *Inj Epidemiol* 2015;2:31.
38 Weiss NS. What findings are needed to advocate personalized (precision) prevention of disease? *Am J Public Health* 2017;107:86–7.
39 Hernán MA RJ. Causal inference. 2017. Updated https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/
40 Mansournia MA, Altman DG. Invited commentary: methodological issues in the design and analysis of randomised trials. *Br J Sports Med* 2018;52:553–5.
41 Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. *Eur J Epidemiol* 2015;30:1101–10.
42 Mansournia MA, Higgins JP, Sterne JA, *et al*. Biases in randomized trials: A conversation between trialists and epidemiologists. *Epidemiology* 2017;28:54–9.
43 Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol* 2013;42:860–9.
44 Shakiba M, Mansournia MA, Salari A, *et al*. Accounting for time-varying confounding in the relation between obesity and coronary heart disease: Analysis with G-estimation, the atherosclerosis risk in communities (ARIC) study. *Am J Epidemiol*. In Press. 2017.
45 Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189.
46 Mansournia MA, Danaei G, Forouzanfar MH, *et al*. Effect of physical activity on functional performance and knee pain in patients with osteoarthritis : analysis with marginal structural models. *Epidemiology* 2012;23:631–40.
47 Gharibzadeh S, Mohammad K, Rahimiforoushani A, *et al*. Standardization as a tool for causal inference in medical research. *Arch Iran Med* 2016;19:666–70.
48 Casals M, Finch CF. Sports Biostatistician: a critical member of all sports science and medicine teams for injury prevention. *Inj Prev* 2017;23:423–7.
49 Gabbett TJ. Influence of training and match intensity on injuries in rugby league. *J Sports Sci* 2004;22:409–17.
50 Gabbett TJ. Reductions in pre-season training loads reduce training injury rates in rugby league players. *Br J Sports Med* 2004;38:743–9.
51 Gabbett TJ, Domrow N. Risk factors for injury in subelite rugby league players. *Am J Sports Med* 2005;33:428–34.
52 Gabbett TJ, Domrow N. Relationships between training load, injury, and fitness in sub-elite collision sport athletes. *J Sports Sci* 2007;25:1507–19.
53 Brooks JH, Fuller CW, Kemp SP, *et al*. An assessment of training volume in professional rugby union and its impact on the incidence, severity, and nature of match and training injuries. *J Sports Sci* 2008;26:863–73.
54 Gabbett TJ. The development and application of an injury prediction model for noncontact, soft-tissue injuries in elite collision sport athletes. *J Strength Cond Res* 2010;24:2593–603.
55 Killen NM, Gabbett TJ, Jenkins DG. Training loads and incidence of injury during the preseason in professional rugby league players. *J Strength Cond Res* 2010;24:2079–84.
56 Gabbett TJ, Jenkins DG. Relationship between training load and injury in professional rugby league players. *J Sci Med Sport* 2011;14:204–9.
57 Hulin BT, Gabbett TJ, Lawson DW, *et al*. The acute:chronic workload ratio predicts injury: high chronic workload may decrease injury risk in elite rugby league players. *Br J Sports Med* 2016;50:231–6.
58 Cross MJ, Williams S, Trewartha G, *et al*. The influence of in-season training loads on injury risk in professional rugby union. *Int J Sports Physiol Perform* 2016;11:350–5.
59 Dennis R, Farhart P, Goumas C, *et al*. Bowling workload and the risk of injury in elite cricket fast bowlers. *J Sci Med Sport* 2003;6:359–67.
60 Dennis RJ, Finch CF, Farhart PJ. Is bowling workload a risk factor for injury to Australian junior cricket fast bowlers? *Br J Sports Med* 2005;39:843–6.
61 Orchard JW, James T, Portus M, *et al*. Fast bowlers in cricket demonstrate up to 3- to 4-week delay between high workloads and increased risk of injury. *Am J Sports Med* 2009;37:1186–92.
62 Hulin BT, Gabbett TJ, Blanch P, *et al*. Spikes in acute workload are associated with increased injury risk in elite cricket fast bowlers. *Br J Sports Med* 2014;48:708–12.
63 Orchard JW, Blanch P, Paoloni J, *et al*. Cricket fast bowling workload patterns as risk factors for tendon, muscle, bone and joint injuries. *Br J Sports Med* 2015;49:1064–8.
64 Lovell G, Galloway H, Hopkins W, *et al*. Osteitis pubis and assessment of bone marrow edema at the pubic symphysis with MRI in an elite junior male soccer squad. *Clin J Sport Med* 2006;16:117–22.
65 Brink MS, Visscher C, Arends S, *et al*. Monitoring stress and recovery: new insights for the prevention of injuries and illnesses in elite youth soccer players. *Br J Sports Med* 2010;44:809–15.
66 Rogalski B, Dawson B, Heasman J, *et al*. Training and game loads and injury risk in elite Australian footballers. *J Sci Med Sport* 2013;16:499–503.
67 Colby M, Dawson B, Heasman J, *et al*. Training and game loads and injury risk in elite australian footballers. *J Strength Cond Res* 2014.
68 Lyman S, Fleisig GS, Waterbor JW, *et al*. Longitudinal study of elbow and shoulder pain in youth baseball pitchers. *Med Sci Sports Exerc* 2001;33:1803–10.
69 Lyman S, Fleisig GS, Andrews JR, *et al*. Effect of pitch type, pitch count, and pitching mechanics on risk of elbow and shoulder pain in youth baseball pitchers. *Am J Sports Med* 2002;30:463–8.
70 Anderson L, Triplett-McBride T, Foster C, *et al*. Impact of training patterns on incidence of illness and injury during a women's collegiate basketball season. *J Strength Cond Res* 2003;17:734–8.
71 Wilson F, Gissane C, Gormley J, *et al*. A 12-month prospective cohort study of injury in international rowers. *Br J Sports Med* 2010;44:207–14.
72 Visnes H, Bahr R. Training volume and body composition as risk factors for developing jumper's knee among young elite volleyball players. *Scand J Med Sci Sports* 2013;23:607–13.
73 Wheeler K, Kefford T, Mosler A, *et al*. The volume of goal shooting during training can predict shoulder soreness in elite female water polo players. *J Sci Med Sport* 2013;16:255–8.
74 Bahr MA, Bahr R. Jump frequency may contribute to risk of jumper's knee: a study of interindividual and sex differences in a total of 11,943 jumps video recorded during training and matches in young elite volleyball players. *Br J Sports Med* 2014;48:1322–6.
75 Hellard P, Avalos M, Guimaraes F, *et al*. Training-related risk of common illnesses in elite swimmers over a 4-yr period. *Med Sci Sports Exerc* 2015;47:698–707.