


# Injury incidence in male elite youth football players is associated with preceding levels and changes in training load

Tania Nilsson <sup>1</sup>, Mats Börjesson,<sup>2,3</sup> Matilda Lundblad,<sup>1,4</sup> Andreas Ivarsson,<sup>5</sup> Dan Fransson<sup>1</sup>

**To cite:** Nilsson T, Börjesson M, Lundblad M, *et al*. Injury incidence in male elite youth football players is associated with preceding levels and changes in training load. *BMJ Open Sport & Exercise Medicine* 2023;**9**:e001638. doi:10.1136/bmjsem-2023-001638

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjsem-2023-001638>).

Accepted 18 September 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

<sup>1</sup>Department of Food and Nutrition and Sports Science, University of Gothenburg, Gothenburg, Sweden

<sup>2</sup>Department of Molecular and Clinical Medicine, University of Gothenburg, Gothenburg, Västra Götaland, Sweden

<sup>3</sup>Department of Medicine, Geriatrics, Acute Care, Sahlgrenska University Hospital, Gothenburg, Sweden

<sup>4</sup>Department of Orthopaedics, Sahlgrenska Academy, Gothenburg, Sweden

<sup>5</sup>School of Health and Welfare, Halmstad University, Halmstad, Sweden

**Correspondence to**  
Tania Nilsson;  
tania.nilsson@gu.se

## ABSTRACT

**Objectives** Elite youth football players miss out on a large part of seasonal training due to injury. Limited research suggests an association between external and internal training load (TL) and injury incidence in elite youth football. This study analysed external and internal TL variables and their association with injury incidence in a group of male elite youth football players over four seasons.

**Methods** Measures of external and internal TL and injury incidence of 56 male elite youth football players (age 17–19 years) were collected throughout four seasons. Heart rate, session rating of perceived exertion and Global Positioning System (GPS) variables were analysed. Individual players' TL during the 30 days leading up to injury was compared with 30-day injury-free control periods. Change in TL through the periods was also analysed.

**Results** Eighty-five injuries were included for analysis, showing that for most TL variables, the average levels were significantly lower during the period leading up to injury. Significant increases for the majority of TL variables were also found during the periods leading up to injury, while the control periods did not show any significant change.

**Conclusion** A lower and/or increasing average TL volume over 30 days might increase the risk of injury in male elite youth football players. Avoiding long-term drops in TL and balance increases in TL might be beneficial to reduce injury risk.

## INTRODUCTION

Injuries affect the youth football player's development, health and possibly burden club/parental finances and healthcare systems.<sup>1</sup> One potential modifiable risk factor for injuries in senior football is rapid increases in external training load (TL) (ie, the physical work to which the athlete is exposed<sup>2</sup> such as distance performed at different speeds),<sup>3</sup> and internal TL (ie, the physiological and psychological response to external<sup>2</sup> TL such as session rate of perceived exertion (S-RPE)).<sup>4</sup> Worth noting is that the majority

## WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Injuries are common in youth football.
- ⇒ The association between training load (TL) and injuries in youth football is currently not established.

## WHAT THIS STUDY ADDS

- ⇒ A low and/or increasing average TL volume during 30 days may be associated with injury risk in male elite youth football players.
- ⇒ The average volume of intense accelerations over >2 m/s<sup>2</sup> during 30 days may be associated with injury in male elite youth football players rather than accelerations at lower speeds.
- ⇒ Decelerations may be associated with injury incidence in male elite youth football players.
- ⇒ Monitoring running at higher speeds over 24.8 km/hour might be of higher interest for youth football teams compared with running at lower speeds

## HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ The results of this study indicate that periods with low and/or increasing average TL volume can be periods when players have an increased risk of injury. A well-developed training exposure and a balanced progression of TL could potentially reduce the injury risk.

of the studies conducted on TL and injuries have shown methodological problems, and no study has yet been able to prove a causal relationship between TL and injuries.<sup>5,6</sup>

The association between external TL and injury risk in elite youth football has been studied by Bacon *et al*<sup>7</sup> who found that weekly averages of total distance (TD) but not high-speed running (HSR) were associated with higher overuse injury risk. However, the study did not clarify which speed zone HSR represented. Only 16 injuries were included, and they used a small group of participants (n=41).<sup>7</sup> The association between internal TL and injury risk in elite youth football has been studied by Brink *et al*<sup>8</sup> who found that S-RPE

was associated with injuries. Still, other studies found no association.<sup>5 9 10</sup> Heart rate (HR) during exercise and its association with injury in youth football players is, to our knowledge, lacking in the scientific literature.

The acute:chronic workload ratio (ACWR) has been used in previous studies to calculate correlations between TL and injuries,<sup>11 12</sup> but this research has been questioned due to methodological errors such as the risk of type I errors.<sup>6</sup> The available evidence exploring TL and its association with injury risk is unreliable for providing practical recommendations on using TL to reduce injury risk in youth football players.<sup>6</sup> Therefore, this study aimed to enhance the understanding and contribute to the research field by focusing on TL and its potential association with injury risk. The study aimed to achieve this by examining the absolute values and capturing changes in external and internal TL variables over a 30-day period and their potential association with injury incidence, in a group of Swedish male elite youth football players.

## METHODS

### Participants

The study was designed as an observational cohort study. Data were collected from 56 male elite youth football players aged 17–19 years from one football club in Sweden's first division for this age group (stature: 181.0±7.8 cm, body mass: 71.2±7.2 kg) during the seasons 2018–2021. Baseline data were collected at player inclusion every season. All playing positions (23 defenders, 29 midfielders, 4 attackers) were represented except goalkeepers due to their different nature of activity. Two players (4 %), 17 players (30 %) and 37 players (66 %) participated in three, two and one seasons, respectively. For a power calculation, we decided that an effect of 0.2 for the relationship between the type of sequence and the level of the outcome variable would be of practical interest.<sup>13</sup> The power calculation, using the *simr* package in *r* (V.4.3.1), showed that a minimum of 40 participants was needed to achieve sufficient power (0.80) for the parameter of interest.

### Training load quantification

The following variables were collected and analysed: average HR and S-RPE, total distance (TD: >0 km/hour), high-intensity running distance (HIRd: 15–19.79 km/hour), high-speed running distance (HSRd: 19.8–24.79 km/hour), sprint distance (Sd: 24.8–29.79 km/hour), maximal sprint distance (Msd: >29.8 km/hour), accelerations (Acc: >0.50 m/s<sup>2</sup>), intense accelerations (IA: >2.00 m/s<sup>2</sup>), very intense accelerations (VIA: >3.00 m/s<sup>2</sup>), decelerations (Dec: <−0.50 m/s<sup>2</sup>), intense decelerations (ID: <−2.00 m/s<sup>2</sup>) and very intense decelerations (<−3.00 m/s<sup>2</sup>). The variables defined above were selected due to their relevance to football TL and potential injury.<sup>3 14</sup>

GPS-derived TL and average HR were quantified during all outfield training sessions and games using 10 Hz GPS and heart rate units (Polar Team Pro; Polar

Electro, Kempele, Finland). This specific unit has a high to extremely high accuracy (ICC=0.63–0.99) in different speed parameters and an interunit reliability of 0.62–0.99 in ICC values.<sup>15</sup> The heart rate measure has been validated and used in previous research.<sup>16 17</sup> The unit was placed on a chest strap, and the same player wore the same unit at all sessions and games. The time for minimum effort duration was set to 1 s, and data sets were verified for the number of satellites connected (mean >11) and horizontal dilution of precision (mean <1.2) before being included in the analysis.<sup>18</sup> Rating of perceived exertion (RPE) was collected after every training session (including gym sessions) and games by asking the players: 'How was your workout?'. The players answered on a modified category scale from 0 min to 10 30 min after the session. S-RPE was calculated by multiplying the RPE value by the duration of each training session or game.<sup>19</sup>

All TL variables were analysed by dividing individual players' data into a 30 days leading up to injury, including the day of injury, and comparing this to a 30-day control period. This model was chosen to compare the individual players with themselves, enabling an individual aspect through the statistical analysis. The control period was 30 days without any absence from training due to injury, illness or vacation. The control periods were spread throughout the year, including preseasons and competitive seasons. In cases of missing data (ie, the player forgot the GPS, GPS was out of function, or S-RPE was not collected), an average based on the individual player's game or training values throughout the season was used. This approach has been used in previous research.<sup>4 20</sup>

### Injury definition

The team's medical staff recorded player exposure and time-loss injuries (ie, when a player was absent from future football participation)<sup>21</sup> using a standard exposure and injury form. Players with an existing injury at the start of the study were included in the study, but their particular injury and the exposure details were not included in the injury statistics until the player could take part in full training and game.<sup>21</sup>

### Statistical analysis

Results are presented in figures and tables. Descriptive statistics are presented as mean and SD. Injury incidence was calculated as the number of injuries per 1000 hours of exposure in training and/or games, as is standard.<sup>21</sup>

Due to the nested data, we applied a three-level model specification (eg, see *Mplus* syntax in online supplemental file 1). Within all analyses, days (level 1) were nested in 30-day periods (level 2). The periods were, in turn, nested in individuals (level 3). To analyse if there was any association between the type of periods (ie, the period leading up to an injury with a control period) and the 30-day average in the different TL variables, we included a dichotomous covariate on level 2 for all the analyses. To analyse if there were any changes in the

internal and external TL variables during the periods, the variable days (ie, ranging from 30 to 0) were regressed on the TL variables on level 1.

The When-to-Worry-and-How-to-Avoid-the-Misuse-of-Bayesian-Statistics (WAMBS) checklist was used to guide the report of the analytical strategy and the results. All statistical analyses were estimated using Bayesian estimation in Mplus V.8.0. One of the advantages of the Bayesian estimator is, in comparison to the frequentist estimator, the increased likelihood of producing reliable estimates with small sample sizes.<sup>22</sup> For the analyses, we performed Markov Chain Monte Carlo simulation procedures with a Gibbs sampler with 100 000 iterations. Four Markov chains were implemented for each of the specified parameters. We assessed chain convergence by applying the Gelman and Rubin convergence diagnostic (this process is described in the Mplus manual). We, however, applied a stricter convergence criterion than the default setting (0.01 instead of 0.05). For all analyses, the Gelman and Rubin diagnostic indicated adequate convergence for each of the four chains. In the next step, we visually inspected the trace plots for each parameter. In this inspection, all four chains showed adequate convergence for each parameter. To validate that convergence was obtained and ruled out potential errors related to local convergence, we ran the analyses again using 200 000 interactions. The Gelman and Rubin convergence diagnosis indicated convergence for all parameters in the analyses. Because of the novel design, we decided to rely on default prior settings ( $N(0, 10^{10})$ ) of the software for the variance on all parameters (ie, intercepts, factor loadings and regression).

Model fit was, for all models, assessed using the posterior predictive  $p$  (PP $p$ ) value and it is accompanying 95% CI.<sup>23</sup> We estimated credibility intervals (CI) for all variables within the models. We rejected the null hypothesis if the 95% CI did not include zero,<sup>24</sup> and then considered the results as statistically significant.

### Patient and public involvement statement

Players were involved from the start of the measures. The players participated in their own team's normal training activities each week. The participating clubs will receive a written report on the study results.

## RESULTS

### Exposure and incidence of injury

The players completed, on average, 130 individual football sessions and 31 games per year and performed one strength session per week. Missing data represented 3% of the GPS variables, 3% of the HR variable and 2% of the S-RPE variable. Given the small amount of missing data, the impact was considered negligible. Ninety-five injuries were recorded during the four seasons. The incidence of injury was estimated to be 4.2 injuries per 1000 exposure hours and 12.2 and 3.0 injuries per 1000 exposure hours for games and training, respectively. Traumatic injuries (injury with sudden onset and known cause<sup>25</sup> accounted

**Table 1** Average levels for TL variables during the control periods compared with the periods leading up to injury for 56 male elite youth football players aged 17–19

Variable	B (95% CI)	Control period M (SD)	The period leading up to the injury M (SD)
S-RPE (au)	0.26 (–0.49 to 0.2)	351.04 (12.33)	326.19 (9.76)
Average HR (bpm)	0.28 (–0.49 to –0.05)	71.73 (3.11)	64.38 (2.25)
TD (m)	0.24 (–0.44 to –0.03)	2851.3 (134.7)	2523.3 (107.3)
ACC (ne)	0.23 (–0.43 to –0.02)	312.45 (15.22)	278.77 (11.79)
IA (ne)	0.28 (–0.51 to –0.04)	34.42 (1.81)	30.00 (1.54)
VIA (ne)	0.34 (–0.69 to –0.02)	6.35 (0.54)	5.40 (0.36)
DEC (ne)	0.24 (–0.44 to –0.04)	350.33 (16.80)	309.82 (12.99)
ID (ne)	0.22 (–0.56 to .04)	32.84 (2.21)	28.73 (1.70)
VID (ne)	0.32 (–0.54 to –0.08)	7.08 (0.49)	5.96 (0.39)
HIRd (m)	0.21 (–0.44 to .03)	237.72 (12.38)	212.53 (12.19)
HSRd (m)	0.21 (–0.44 to .03)	105.57 (6.81)	93.13 (6.04)
Sd (m)	0.28 (–0.52 to –0.02)	35.85 (3.12)	29.46 (2.78)
Msd (m)	0.31 (–0.55, –0.07)	4.20 (0.75)	3.06 (0.42)

Acc, accelerations (>0.5 m/s<sup>2</sup>); au, arbitrary units; B, unstandardised beta coefficient; bpm, beats per minute; CI, credibility interval; Dec, decelerations (< –0); HIRd, high-intensity running distance (15–19.79 km/hour); HSRd, high-speed running distance (19.8–24.79 km/h); IA, Intense accelerations (>2 m/s<sup>2</sup>); ID, intense decelerations (<–2 m/s<sup>2</sup>); m, metres; Msd, maximal sprint distance (>29.8 km/h); ne, number of efforts; PPp, posterior predictive p-value; Sd, Sprint distance (24.8–29.9 km/h); S-RPE, session rate of perceived exertion; TD, total distance (>0 km/h); VIA, very intense accelerations (>3 m/s<sup>2</sup>); VID, very intense decelerations (<–3 m/s<sup>2</sup>).

for 71% (n=60) and overuse injuries (injury with insidious onset and no known trauma)<sup>25</sup> represented 29% (n=25) of all injuries. Ten injuries were excluded from further analysis due to a major lack of TL data. In summary, 85 injuries were used for further analysis.

### TL and injury

All models specified to analyse the association between the periods, leading up to injury and the control periods and average level on the TL variables, showed a good fit to the data. The three-level analyses showed significant negative associations between the periods leading up to injury and all TL variables except ID, HIRd and HSRd. More specifically, the average level on all variables was consistently lower during the periods, leading up to injury compared with the control periods (table 1). For all parameter estimates, see online supplemental file 2.

All models specified to analyse if a change in the TL variables during the 30-day periods fits the data well. An average of 20.6 and 19.3 training sessions was conducted during the control periods and periods leading up to injury, respectively. Significant positive changes on all TL variables, except Acc, indicated an increase in TL during the periods, leading up to injury (table 2). There was no significant change in any TL variable during the control

**Table 2** Average change estimates (with 95% CI) for the variables in the control periods and the periods leading up to injury for 56 male elite youth football players aged 17–19 years old

Variable	Change over the 30 days leading up to injury (B, 95% CI)	Changes over the 30 days control period (B, 95% CI)
S-RPE (au)	2.30 (0.97 to 3.62)	0.90 (2.66 to 0.80)
Average HR (bpm)	0.81 (0.52 to 1.09)	0.22 (0.60 to 0.15)
TD (m)	0.38 (0.24 to 0.51)	0.19 (0.16 to 0.20)
ACC (ne)	3.67 (2.28 to 5.03)	0.42 (2.25 to 1.35)
IA (ne)	0.31 (0.13 to 0.51)	0.03 (0.28 to 0.21)
VIA (ne)	0.04 (0.01 to 0.08)	0.01 (0.09 to 0.06)
DEC (ne)	4.03 (2.52 to 5.52)	0.64 (2.65 to 1.31)
ID (ne)	0.26 (0.02 to 0.50)	0.01 (0.35 to 0.36)
VID (ne)	0.09 (0.05 to 0.13)	0.02 (0.04 to 0.08)
HIRd (m)	3.21 (1.74 to 4.67)	0.61 (1.38 to 2.53)
HSRd (m)	1.26 (0.58 to 1.95)	0.32 (0.57 to 1.19)
Sd (m)	0.61 (0.32 to 0.90)	0.02 (0.41 to 0.36)
Msd (m)	0.09 (0.04 to 0.14)	0.02 (0.10 to 0.05)

Acc, accelerations (>0); au, arbitrary units; B, unstandardised beta coefficient; bpm, beats per minute; CI, credibility interval; Dec, decelerations (<-0); HIRd, high-intensity running distance (15–19.79 km/hour); HSRd, high-speed running distance (19.8–24.79 km/h); IA, intense accelerations (>2 m/s<sup>2</sup>); ID, intense decelerations (<-2 m/s<sup>2</sup>); m, metres; Msd, maximal sprint distance (>29.8 km/h); ne, number of efforts; Sd, Sprint distance (24.8–29.9 km/h); S-RPE, session rate of perceived exertion; TD, total distance (>0 km/h); VIA, very intense accelerations (>3 m/s<sup>2</sup>); VID, very intense decelerations (<-3 m/s<sup>2</sup>).

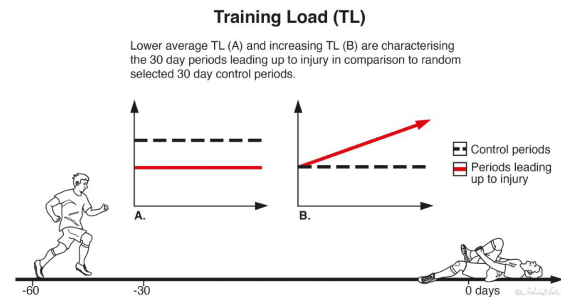
periods. For all parameter estimates, see online supplemental file 2.

## DISCUSSION

The main findings of this study were that external and internal TL is associated with injury incidence in male elite youth football. The average level of all variables, except intense ID, HIRd and HSRd, were significantly lower during the periods leading up to injury in comparison to the control periods. In addition, significant increases in all TL variables except accelerations (Acc) were found during the periods leading up to injury, while no significant change in any TL variable was found during the control periods. In summary, the results of the present study indicate that periods with low and/or increasing average TL volume might be periods where players are at a higher risk of sustaining an injury.

### Injury incidence

Injury incidence was estimated to be 4.2 injuries per 1000 hour of exposure, considerably lower compared with previous findings on elite football players of the same age.<sup>9 10</sup> Injury incidence can be affected by factors



**Figure 1** Association between training load and subsequent injury.

such as methodological differences in injury definition or reporting systems,<sup>26</sup> levels of play, player age and previous injury.<sup>3 27</sup>

### TL and injury

Most TL variables had a lower average volume during the periods leading up to injury, while the control periods had a higher average volume (figure 1). The number of training sessions was relatively equal between the compared periods. Potentially, a high average TL volume over time (as during the control periods) in combination with adequate recovery might increase the strength of the tissues, possibly leading to a better tolerance against increasing TL and subsequent injury.<sup>9 28</sup> An excessive overload or a rapid progression in TL can on the other hand lead to suboptimal adaptations where the strength of the tissues involved might be exceeded, potentially increasing the risk of injury or causing tissue damage.<sup>29 30</sup> Interestingly, the majority of TL variables in this study showed an increase in average volume during the periods leading up to injury (figure 1).

Running at higher speeds puts higher physiological stress on the body than running at lower speeds.<sup>28</sup> In this study, the average running volume at lower speeds under 24.79 km/hour (HIRd and HSRd) was equal between control periods and periods leading up to injury. However, for running speeds over 24.8 km/hour (Sd and Msd), the average volume was lower during the periods leading up to injury. In a recent paper, Edouard *et al*<sup>31</sup> discussed that optimal exposure to maximal or near-maximal running velocity might be a protective factor for hamstring injuries and that a lack of regular preparatory sprint training might lead to a higher risk of sprint-related injuries. The authors suggested that athletes should be exposed to sprinting regularly to avoid a long-term drop in sprinting volume and potential injury.<sup>31</sup> This approach can be adapted to the present study results as well. Monitoring running at higher speeds might, therefore, interest youth football teams.

The results from the present study also indicate that changes in volume in intense accelerations (IA and VIA) over >2 m/s<sup>2</sup> should be of particular interest to monitor rather than accelerations at lower speeds, since the variable Acc that includes accelerations in

lower speeds was the only variable not showing an association between change in volume during 30 days and injury incidence. Accelerating at higher speeds puts higher physiological demands on the body than accelerating at lower speeds,<sup>32</sup> which may be why increasing volumes of accelerations at higher speeds is associated with injury in this study.

Dec have been associated with injury risk in senior football,<sup>14</sup> and future research has been requested to evaluate its association with injury in youth football.<sup>33</sup> Our results show an increasing average volume in Dec during the periods leading up to injury compared with the control periods, which indicates that Dec is associated with injury risk in elite youth football players.

Previous literature investigating the association between S-RPE and risk of injury for youth football players shows low support when using the acute-chronic workload ratio.<sup>5 9 10</sup> HR has previously been associated with injuries in elite senior football<sup>34</sup> but not in elite youth football. In the present study, S-RPE and average HR were associated with injury risk since both variables had a lower and increasing average volume during the periods leading up to injury compared with the control periods.

Future multifactorial research of high quality is required to further assess the relationships between measures of TL and injury risk for football players. Researchers should also consider examining injuries and changes over time in TL for shorter periods than 30 days.

### Clinical implications

Evaluating TL from an individual perspective can identify periods with a low and/or increasing average TL volume where players might have an increased risk of injury. A well-developed training exposure and a balanced progression of TL can be beneficial to reduce injury risk. To deal with this, every athlete should be evaluated individually and monitored regularly, and the data should be compared longitudinally.

### Limitations

We acknowledge that there are major methodological challenges in conducting this type of research.<sup>6 35</sup> Several previous studies have used the ACWR to calculate correlations between TL and injuries and the possibilities of predicting injuries.<sup>11 12</sup> These studies have been questioned due to methodological errors,<sup>5 15</sup> such as the risk of type I errors due to calculation associations by using multiple combinations.<sup>6</sup> Other methodological errors considered are that the ACWR is used to control for chronic workload (the denominator variable), which is assumed to influence the acute workload (the numerator). Still, researchers have shown that the ACWR fails to normalise the numerator for the denominator.<sup>36</sup> Impellizzeri and colleagues describe this problem well in the study

Training load and Its Role in Injury Prevention, Part 2: Conceptual and Methodologic Pitfalls<sup>6</sup>: *This failure adds unnecessary “noise,” increases the risk of artifact and makes the results difficult to interpret, as the practitioner cannot determine whether the acute or chronic TL is driving the ACWR.*<sup>6</sup> In the present study, we wanted to investigate TL's association with injuries by not using a ratio but instead using a multilevel model recommended by Windt *et al.*<sup>37</sup> We investigated absolute values and captured changes in TL over 30 days. Still, we did not look at cumulative values or day-to-day fluctuation in TL, which can be considered a limitation of the study. Recent studies have also suggested that non-linear models might be beneficial in discovering relationships between TL and injury.<sup>35</sup> A more consistent methodology for TL and injury risk studies can enable better comparability and reproducibility within this research area in the future.

Confounding factors can affect whether a player gets injured, such as the possibility and quality of recovery between training sessions,<sup>29</sup> the implementation of injury prevention programmes<sup>6</sup> or previous injury.<sup>27</sup> This information was unavailable and, therefore, not included in the analysis in this study. Also, we did not have enough data to perform separate analyses for traumatic and overuse injuries. This is a limitation given the potential risk factors for these two types of injuries.

Sample size and data from only one club can be a study limitation. Another limitation is that the club recorded the data, and it is unknown whether the practitioners adapted TL to protect players. We acknowledge that the diversity was limited to this population, excluding male youth football players at grassroots levels and those from other geographical regions in Sweden.

### CONCLUSION

Injury is a complex and dynamic outcome affected by many different risk factors, often with no predictable pattern,<sup>38</sup> which makes the goal of decreasing injury risk hard to reach. We cannot prove a causal relationship between TL and injury. Still, this study suggests that periods with low and/or increasing average TL volume might be periods when players have an increased risk of injury. Therefore, evaluating TL from an individual perspective can identify periods when players potentially have an increased risk of injury, representing one piece of the injury prevention puzzle.<sup>39</sup> Practitioners may use these indicators in TL combined with their clinical experience and knowledge of physiology and basic training principles to decrease injury risk.

**Twitter** Tania Nilsson @NilssonTania

**Acknowledgements** The authors would like to thank Mehdi Hamdani and Katrin Muller for their input and help during the processing of the TL variables and

Karin Lodin for creating the figure describing the results of the study. Also, much gratitude to Victor Stoltz and Torstein Dalen-Lorentsen for valuable feedback on the draft manuscript.

**Contributors** All authors listed met the conditions required for full authorship. TN and DF designed the initial study proposal, which was then completed together with MB, ML and AI. From there, all five authors met regularly to ensure that a scientifically sound study design was created. TN collated the data and then processed the data with help of DF and AI. AI advised and checked the statistical analysis. All authors contributed to the interpretation of data and its presentation. TN wrote the main body of the article, which was revised multiple times by all authors before being approved for submission. MB is responsible for the overall content as guarantor.

**Funding** The study was funded by grants from Folksam Research Council and IF Elfsborg.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by Swedish Ethical Review Authority. Reference number 2020-00876 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 upon reasonable request. Except for the data presented in the study, more information about injury descriptions or raw data are the property of the club and therefore only available to the researchers, the players and the club only.

**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 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID ID

Tania Nilsson <http://orcid.org/0009-0001-2715-3133>

#### REFERENCES

- Jones S, Almousa S, Gibb A, *et al*. Injury incidence, prevalence and severity in high-level male youth football: A systematic review. *Sports Med* 2019;49:1901.
- Halson SL. Monitoring training load to understand fatigue in athletes. *Sports Med* 2014;44 Suppl 2(Suppl 2):S139–47.
- Malone S, Owen A, Mendes B, *et al*. High-speed running and Sprinting as an injury risk factor in soccer: can well-developed physical qualities reduce the risk? *J Sci Med Sport* 2018;21:257–62.
- McCall A, Dupont G, Ekstrand J. Internal workload and non-contact injury: a one-season study of five teams from the UEFA elite club injury study. *Br J Sports Med* 2018;52:1517–22.
- 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.
- Impellizzeri FM, McCall A, Ward P, *et al*. Training load and its role in injury prevention, part 2: conceptual and Methodologic pitfalls. *J Athl Train* 2020;55:893–901.
- Bacon CS, Mauger AR. Prediction of Overuse injuries in professional U18-U21 Footballers using Metrics of training distance and intensity. *J Strength Cond Res* 2017;31:3067–76.
- 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.
- Raya-González J, Nakamura FY, Castillo D, *et al*. Determining the relationship between internal load markers and Noncontact injuries in young elite soccer players. *Int J Sports Physiol Perform* 2019;14:421–5.
- Delecroix B, Delavaud B, Dawson B, *et al*. Workload and injury incidence in elite football Academy players. *J Sports Sci* 2019;37:2768–73.
- Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter and harder? *Br J Sports Med* 2016;50:273–80.
- 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.
- Funder DC, Ozer DJ. Evaluating effect size in psychological research: sense and nonsense. *Advances in Methods and Practices in Psychological Science* 2019;2:156–68.
- Jaspers A, Kuyvenhoven JP, Staes F, *et al*. Examination of the external and internal load indicators' association with Overuse injuries in professional soccer players. *J Sci Med Sport* 2018;21:579–85.
- Akyildiz Z, Yildiz M, Clemente FM. The Reliability and accuracy of polar team pro GPS units. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology* 2022;236:83–9.
- Hernando D, Garatachea N, Almeida R, *et al*. Validation of heart rate monitor polar Rs800 for heart rate variability analysis during exercise. *J Strength Cond Res* 2018;32:716–25.
- Gillinov S, Etiwy M, Wang R, *et al*. Variable accuracy of Wearable heart rate monitors during aerobic exercise. *Med Sci Sports Exerc* 2017;49:1697–703.
- Vickery WM, Dascombe BJ, Baker JD, *et al*. Accuracy and reliability of GPS devices for measurement of sports-specific movement patterns related to cricket, tennis, and field-based team sports. *J Strength Cond Res* 2014;28:1697–705.
- Foster C, Florhaug JA, Franklin J, *et al*. A new approach to monitoring exercise training. *J Strength Cond Res* 2001;15:109–15.
- Ehrmann FE, Duncan CS, Sindhusake D, *et al*. GPS and injury prevention in professional soccer. *J Strength Cond Res* 2016;30:360–7.
- Fuller CW, Ekstrand J, Junge A, *et al*. Consensus statement on injury definitions and data collection procedures in studies of football (soccer) injuries. *Br J Sports Med* 2006;40:193–201.
- Song X-Y, Lee S-Y. A Tutorial on the Bayesian approach for analyzing structural equation models. *Journal of Mathematical Psychology* 2012;56:135–48.
- Muthén B, Asparouhov T. Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychol Methods* 2012;17:313–35.
- Zyphur MJ, Oswald FL. Bayesian estimation and inference: A user's guide. *Journal of Management* 2015;41:390–420.
- Ekstrand J, Hägglund M, Waldén M. Injury incidence and injury patterns in professional football: the UEFA injury study. *Br J Sports Med* 2011;45:553–8.
- Bahr R, Clarsen B, Derman W, *et al*. International Olympic committee consensus statement: methods for recording and reporting of Epidemiological data on injury and illness in sport 2020 (including STROBE extension for sport injury and illness surveillance (STROBE-SIIS)). *Br J Sports Med* 2020;54:372–89.
- Hägglund M, Waldén M, Ekstrand J. Exposure and injury risk in Swedish elite football: a comparison between seasons 1982 and 2001. *Scand J Med Sci Sports* 2003;13:364–70.
- Minetti AE, Moia C, Roi GS, *et al*. Energy cost of walking and running at extreme uphill and downhill slopes. *Journal of Applied Physiology* 2002;93:1039–46.
- Kalkhoven JT, Watsford ML, Impellizzeri FM. A conceptual model and detailed framework for stress-related, strain-related, and Overuse athletic injury. *J Sci Med Sport* 2020;23:726–34.
- Edwards WB. Modeling Overuse injuries in sport as a mechanical fatigue phenomenon. *Exerc Sport Sci Rev* 2018;46:224–31.
- Edouard P, Mendiguchia J, Guex K, *et al*. Sprinting: a key piece of the hamstring injury risk management puzzle. *Br J Sports Med* 2023;57:4–6.
- Hatamoto Y, Yamada Y, Fujii T, *et al*. A novel method for calculating the energy cost of turning during running. *Open Access J Sports Med* 2013;4(default):117–22.
- McBurnie AJ, Harper DJ, Jones PA, *et al*. Deceleration training in team sports: another potential 'vaccine' for sports-related injury? *sports MED* 2021 *Sports Med* 2022;52:1–12.

- 34 Owen AL, Forsyth JJ, Wong DP, *et al.* Heart rate-based training intensity and its impact on injury incidence among elite-level professional soccer players. *J Strength Cond Res* 2015;29:1705–12.
- 35 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.
- 36 Lolli L, Batterham AM, Hawkins R, *et al.* The acute-to-chronic workload ratio: an inaccurate Scaling index for an unnecessary normalisation process *Br J Sports Med* 2019;53:1510–2.
- 37 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.
- 38 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.
- 39 West SW, Clubb J, Torres-Ronda L, *et al.* More than a metric: how training load is used in elite sport for athlete management. *Int J Sports Med* 2021;42:300–6.

TITLE: Syntax for the three-level analysis with Sequence Type as co-variate on level 2.

DATA:

FILE IS "Study.dat";

VARIABLE:

NAMES ARE ID Clus Type N\_Time !ID = Level 3 Clus = Level 2 Type = Sequence Type

S\_RPE HR TD ACC IA VIA DEC ID VID HIRd

HSRd Sd Msd;

USEVARIABLES ARE

Type

S\_RPE

;

MISSING ARE ALL (-111);

CLUSTER = ID Clus;

BETWEEN (Clus) = Type;

ANALYSIS:

TYPE IS THREELEVEL RANDOM;

ESTIMATOR = BAYES;

Bconvergence = .01;

Biterations = 40000 (10000);

Processor = 4;

Chains = 4;

thin = 10;

BSEED = 1234;

MODEL:

%BETWEEN ID%

%BETWEEN Clus%

S\_RPE ON Type;

%WITHIN%

OUTPUT: STDYX



CINTERVAL TECH1 TECH4;

PLOT:

TYPE IS PLOT2;

TITLE: Syntax for the three-level analysis with time as a co-variate on level 1.

DATA:

FILE IS "Study.dat";

VARIABLE:

NAMES ARE ID Clus Type N\_Time !ID = Level 3 Clus = Level 2 Type = Sequence Type

S\_RPE HR TD ACC IA VIA DEC ID VID HIRd

HSRd Sd Msd;

USEVARIABLES ARE

N\_Time

S\_RPE

;

MISSING ARE ALL (-111);

CLUSTER = ID Clus;

Within = N\_Time;

ANALYSIS:

TYPE IS THREELEVEL RANDOM;

ESTIMATOR = BAYES;

Bconvergence = .01;

Biterations = 40000 (10000);

Processor = 4;

Chains = 4;

thin = 10;

BSEED = 1234;

MODEL:

%BETWEEN ID%

%BETWEEN Clus%

%WITHIN%

S\_RPE ON N\_Time;

OUTPUT: STDYX

CINTERVAL TECH1 TECH4;

PLOT:

TYPE IS PLOT2;

Table 3 Supplementary information

	S-RPE (au)	Average HR (bpm)	TD (m)	ACC (ne)	IA (ne)	VIA (ne)	DEC (ne)	ID (ne)	VID (ne)	HIRd (m)	HSRd (m)	Sd (m)	Msd (m)
<b>Fixed effects level 2</b>													
sequence type	-0.26 [-0.49, -0.02]	-0.28 [-0.49, -0.05]	-0.24 [-0.44, -0.03]	-0.23 [-0.43, -0.02]	-0.28 [-0.51, -0.04]	-0.34 [-0.69, -0.02]	-0.24 [-0.44, -0.04]	-0.25 [-0.56, 0.04]	-0.32 [-0.54, 0.08]	-0.21 [-0.44, 0.03]	-0.21 [-0.44, 0.03]	-0.28 [-0.52, 0.02]	-0.31 [-0.55, -0.07]
<b>Random variances level 1</b>													
Variance	96728.70 [92657.60, 100997.02]	4502.79 [4212.67, 4702.00]	100617.70 [96424.80, 105094.70]	NA [99728.31, 108714.92]	1885.06 [1805.77, 1968.58]	137.63 [131.82, 143.72]	NA [119593.43, 130375.34]	3081.62 [2951.53, 3218.79]	96.72 [92.67, 101.01]	NA [113822.89, 124086.77]	25377.09 [24304.97, 26497.56]	4522.20 [4331.22, 4722.48]	147.45 [141.23, 154.04]
<b>Random variances level 2</b>													
Residual variance	2454.70 [1166.66, 4238.50]	146.39 [75.76, 247.03]	4080.00 [2383.40, 6376.60]	4729.83 [2794.96, 7427.22]	50.80 [24.03, 89.65]	1.68 [0.29, 3.79]	5595.11 [3310.91, 8783.95]	47.74 [12.77, 98.43]	2.82 [1.38, 4.90]	3424.08 [1634.83, 5928.64]	746.48 [364.58, 1308.14]	106.18 [44.62, 197.71]	3.70 [1.62, 6.68]
<b>Random variances level 3</b>													
Variance	897.34 [123.05, 2469.92]	76.74 [15.51, 189.47]	771.30 [49.30, 2703.60]	1570.93 [206.34, 4336.35]	39.85 [11.25, 90.49]	3.45 [1.56, 6.82]	1924.17 [278.44, 5213.49]	50.27 [15.32, 112.55]	3.61 [1.53, 7.26]	1616.93 [196.46, 4524.18]	738.12 [236.47, 1615.29]	163.89 [69.01, 332.75]	6.76 [3.23, 13.06]
<b>Model fit indices</b>													
PPp	0.49	0.48	0.49	0.49	0.49	0.48	0.49	0.48	0.49	0.49	0.48	0.47	0.48
95 % Confidence Interval	[-10.91, 11.99]	[-11.11, 11.84]	[-11.01, 12.17]	[-10.64, 11.72]	[-11.12, 12.03]	[-10.55, 12.07]	[-10.67, 11.73]	[-11.05, 11.64]	[-10.87, 12.00]	[-10.89, 11.69]	[-10.90, 11.16]	[-9.79, 11.56]	[-10.67, 11.94]
PPp: Posterior Predictive p-value, B: Unstandardized Beta Coefficient, CI: Credibility interval, SD: standard deviation, au: arbitrary units, bpm: beats per minute, m: meters, ne: number of efforts, S-RPE: Session rate of perceived exertion, TD: Total distance (>0 km/h), Acc: Accelerations (>0,5 m/s <sup>2</sup> ), IA: Intense accelerations (>2 m/s <sup>2</sup> ), VIA: Very intense accelerations (>3 m/s <sup>2</sup> ), Dec: Decelerations (< -0,5 m/s <sup>2</sup> ), ID: Intense decelerations (< -2 m/s <sup>2</sup> ), VID: Very intense decelerations (< -3 m/s <sup>2</sup> ), HIRd: High-intensity running distance (15 km/h - 19,79 km/h), HSRd: High-speed running distance (19,-24,79 km/h), Sd: Sprint distance (24,8 – 29,9 km/h), Msd: Maximal sprint distance (> 29,8 km/h).													

Table 4 Supplementary information

	S-RPE (au)	Average HR (bpm)	TD (m)	ACC (ne)	IA (ne)	VIA (ne)	DEC (ne)	ID (ne)	VID (ne)	HIRd (m)	HSRd (m)	Sd (m)	Msd (m)
<b>Fixed effects level 1</b>													
time	2.30 [0.97, 3.62]	0.81 [0.52, 1.09]	0.38 [0.24, 0.51]	3.67 [2.28, 5.03]	0.31 [0.13, 0.51]	0.04 [-0.01, 0.08]	4.03 [2.52, 5.52]	0.26 [0.02, 0.50]	0.09 [0.05, 0.13]	3.21 [1.74, 4.67]	1.26 [0.58, 1.95]	0.61 [0.32, 0.90]	0.09 [0.04, 0.14]
<b>Random variances level 1</b>													
Residual variance	97971.35 [92881.33, 103605.96]	4493.19 [4258.53, 4751.90]	99636.30 [94430.20, 105339.90]	NA [98952.13, 110412.74]	1971.01 [1867.98, 2083.74]	124.62 [118.10, 131.77]	NA [117503.89, 131089.09]	3158.97 [2992.68, 3339.63]	90.01 [85.33, 95.13]	NA [111492.28, 124365.02]	25908.32 [24555.48, 27381.14]	4553.49 [4314.40, 4812.91]	120.68 [114.37, 127.56]
<b>Random variances level 2</b>													
Variance	2670.24 [1006.66, 5183.00]	169.06 [77.23, 307.03]	4830.00 [2630.40, 8191.50]	5539.88 [2972.54, 9407.94]	77.88 [35.71, 139.17]	2.82 [0.73, 5.88]	6352.78 [3341.26, 10861.82]	88.76 [35.30, 166.44]	3.92 [1.68, 7.37]	5269.93 [2513.01, 9228.07]	1170.67 [526.95, 2096.72]	121.28 [26.94, 278.93]	2.19 [0.31, 5.44]
<b>Random variances level 3</b>													
Variance	615.40 [27.68, 2533.13]	39.41 [1.61, 154.24]	564.70 [20.40, 2855.50]	990.93 [41.02, 4190.43]	18.32 [0.74, 72.36]	1.32 [0.08, 4.07]	1291.14 [57.11, 5140.47]	16.31 [0.61, 71.43]	2.30 [0.17, 6.55]	1125.49 [39.74, 5182.58]	381.13 [16.18, 1411.12]	151.72 [16.54, 383.91]	3.30 [0.73, 7.63]
<b>Model fit indices</b>													
PPp	0.47	0.48	0.47	0.47	0.48	0.47	0.46	0.47	0.46	0.47	0.48	0.47	0.46
95 % Confidence Interval	[-10.36, 11.39]	[-10.46, 11.80]	[-10.29, 11.69]	[-10.40, 11.19]	[-10.82, 10.98]	[-10.42, 11.34]	[-10.45, 11.78]	[-10.31, 11.73]	[-10.18, 11.33]	[-10.29, 11.80]	[-10.05, 11.10]	[-10.44, 11.75]	[-10.19, 11.58]
PPp: Posterior Predictive p-value, B: Unstandardized Beta Coefficient, CI: Credibility interval, SD: standard deviation, au: arbitrary units, bpm: beats per minute, m: meters, ne: number of efforts, S-RPE: Session rate of perceived exertion, TD: Total distance (>0 km/h), Acc: Accelerations (>0,5 m/s <sup>2</sup> ), IA: Intense accelerations (>2 m/s <sup>2</sup> ), VIA: Very intense accelerations (>3 m/s <sup>2</sup> ), Dec: Decelerations (< -0,5 m/s <sup>2</sup> ), ID: Intense decelerations (< -2 m/s <sup>2</sup> ), VID: Very intense decelerations (< -3 m/s <sup>2</sup> ), HIRd: High-intensity running distance (15 km/h - 19,79 km/h), HSRd: High-speed running distance (19,-24,79 km/h), Sd: Sprint distance (24,8 - 29,9 km/h), Msd: Maximal sprint distance (> 29,8 km/h).													