

# Assessing the relationship between training load and injury in ultramarathon runners: a novel approach using Generalised Additive Models

TL Burgess,<sup>1,2</sup>  PhD; P Durand,<sup>1</sup>  MPhil; K Buchholtz,<sup>3,4</sup>  PhD

<sup>1</sup> Division of Physiotherapy, Faculty of Health Sciences, University of Cape Town

<sup>2</sup> Centre for Medical Ethics and Law, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

<sup>3</sup> Department of Health, LUNEX University of Applied Sciences, Luxembourg

<sup>4</sup> Luxembourg Health & Sport Sciences Research Institute A.s.b.l., 50, Avenue u Parc des Sports, 4671 Oberkorn Differdange, Luxembourg

Corresponding author: K Buchholtz (kbuchholtz@lunex.lu)

**Background:** Ultramarathon running presents significant injury risks, and monitoring training loads may identify risk factors for injury. Injury surveillance studies are required to better assess injury prevalence and its relationship to training loads.

**Objectives:** To determine the incidence and nature of running-related injuries and associated training loads in runners 12 weeks before and two weeks after the 2018 Comrades ultramarathon.

**Methods:** One hundred and six participants were recruited. Their weekly injury and training load data (distance, duration, frequency and acute-chronic workload ratio) were obtained retrospectively over 14 weeks. The relationship between training load variables and injury risk was modelled using Generalised Additive Models.

**Results:** The running-related injury incidence was 8/1000 hours. The overall injury proportion was 40%. The commonly injured structures were muscles (47%) followed by tendons (24%). Commonly reported anatomical areas of injury were the knee (26%) and hip (19%). Lower training load distance in the 12 weeks leading up to the race was linked to a higher risk of injury ( $p=0.02$ ), primarily occurring during or after the race. Weekly training frequency and injury risk showed a significant heterogeneous relationship ( $p=0.02$ ). The effect of the acute to chronic workload ratio on injury risk was minimal ( $p=0.3$ ).

**Conclusion:** Lower training loads were associated with a higher risk for injury, and the frequency of running training per week influenced injury risk. Insufficient training may not prepare the runners for the demands of the ultradistance race. Sudden changes in training load (evident in the acute training load measurements) appeared to have a minimal effect on injury risk. The non-linear relationship between several training load variables and injury risk can successfully be modelled using Generalised Additive Models, which may improve the accuracy of injury prediction modelling in ultramarathon runners.

**Keywords:** running-related injuries, Comrades, acute-chronic workload

*S Afr J Sports Med* 2025;37:1-8. DOI: 10.17159/2078-516X/2025/v37i1a20747



Participation in ultramarathon running has seen exponential growth over recent years.<sup>[1,2]</sup> This may be attributed to the numerous health benefits, ease of access and affordability of the sport.<sup>[3]</sup> Conversely, ultramarathons have also been associated with a high prevalence of running-related injuries (RRIs), resulting in time loss from training and competition and socio-economic costs related to missing work due to injury or medical care-related costs.<sup>[1,4]</sup>

Previous studies have reported the prevalence of RRIs to range between 30% and 80%.<sup>[3,5]</sup> Across 13 studies, an RRI incidence rate of 2.5 to 33 injuries per 1000 hours of running exposure has been reported.<sup>[4]</sup> The wide range of prevalence and incidence are hypothesised to be related to a lack of standardisation between studies in injury recording and reporting methods. To address the lack of standardisation of epidemiological data, the International Olympic Committee (IOC) released a consensus statement with recommendations for researchers when reporting sport-related injuries.<sup>[6]</sup> Ongoing injury surveillance studies on ultramarathon runners that report epidemiological data according to recent recommendations have received global attention.

Risk factors for injury can be broadly categorised as either modifiable or non-modifiable.<sup>[7]</sup> Training load is a well-

known modifiable risk factor for injury, showing positive or negative effects on RRIs.<sup>[8]</sup> Although research aimed at predicting the relationship between training load and injury risk in ultramarathon running specifically is limited, a systematic review on endurance sports reported a significant association between training distance, duration and frequency, and the incidence of injuries.<sup>[9]</sup> Therefore, appropriately modifying and managing various aspects of the training load may reduce the high injury rate reported.<sup>[7]</sup> The acute-to-chronic workload ratio (ACWR) is used globally as a prognostic tool in determining athletic performance, well-being and injury risk.<sup>[10,11]</sup> The ACWR compares athletes' acute training load (one week) with the chronic training load (three to six weeks rolling average) to evaluate athletic preparedness.<sup>[9,11]</sup> The ACWR is calculated as (acute workload [km in the past 7 days])/(chronic workload [average of km per week over three to six weeks]). The researcher determines the number of weeks used for the rolling chronic workload based on the needs of the study.<sup>[11]</sup>

However, the calculation of the ACWR has been debated, with researchers arguing that the mathematical coupling of the chronic and acute load values is questionable.<sup>[10,12]</sup> Wang et al.<sup>[13]</sup> and Wang et al.<sup>[10]</sup> demonstrated that the relationship between the ACWR and injury risk is best represented when using an uncoupled ACWR in Generalised Additive Models.

Recent calls for better handling of non-linear data in sports medicine research prompted this [10,13,14,15]. Generalised Additive Models are automatic flexible statistical methods that may be used to identify and describe non-linear regression effects.<sup>[16]</sup> It has been proposed that modelling training load and injury risk using Generalised Additive Models might be better suited to model the flexible relationship across the diverse exposure range of training load and injury risk.<sup>[10,13]</sup>

The objectives of this study were to determine the incidence and nature (anatomical region, type, and time loss) of running-related injuries and their association with training loads in runners 12 weeks before and two weeks after the 2018 Comrades ultramarathon.

## Methods

### Study design

This study had a retrospective, cross-sectional design with repeated measures. One hundred and six participants were recruited, from whom weekly injury and training load data from 12 weeks before and two weeks after the 2018 Comrades ultramarathon were obtained retrospectively. Training load data were gathered from data stored on Global Positioning System (GPS) watches, and the participants reported injuries retrospectively.

### Setting and participants

Participants were recruited from eight running clubs and three community-organised road races situated in Pretoria, Johannesburg, and Midrand regions in South Africa. The Comrades ultramarathon is the largest in the world, providing a substantial platform for conducting a surveillance study on ultramarathon participants. In 2018, Comrades was a 'down run', and the approximately 90-kilometre (km) route covered a hilly course for the first 60 km, with the last 30 km featuring a predominantly downhill profile, which resulted in the 'down-run' title, in comparison to alternate years in the opposite direction 'up-run'. The total elevation gain during the course is approximately 1095 meters (m), with the highest point at 823 m above sea level.

Both male and female ultramarathon runners older than 20 years (minimum age requirement for the race) were included in this study. Participants were only included if they had completed the 2018 Comrades ultramarathon and used GPS watches to track and store their training data.

### Data capturing and analysis

Two methods of data capturing were used: 'Face-to-face facilitation group,' where the researcher conducted in-person data collection with each participant, or 'Online facilitation,' via Zoom or WhatsApp calls. The participants completed the informed consent form, a baseline questionnaire on training history, and the injury logbook.

Descriptive characteristics of the participants' demographics were described (mean, SD, 95% CI). Participants were allocated to the 'injured' group or 'uninjured' group and analysed using incidence percentages

for the population, injured and non-injured groups (IBM Corp, USA. IBM SPSS Statistics for Macintosh, Version 25.0, 2017).

Training variables (distance, duration, frequency, and ACWR) were presented as cumulative 14-week and weekly averages. The ACWR was calculated by dividing the workload by km per week (acute) and the average of the preceding four weeks (chronic). An ACWR ratio of 1 may be interpreted as the runner running equivalent km in the current week as the average of the previous weeks. In contrast, a ratio below 1 means they have completed less training load during the current week than the previous average load. A ratio above 1 means that the runner completed more acute training load (in that current week) than the average of the previous weeks.<sup>[11]</sup>

An independent samples t-test was conducted to identify any significant differences in the training variables. A Mann-Whitney U test was used to compare the number of races. Cohen's effect sizes were used to quantify the size of the difference. One way Analysis of Variance (ANOVA) and repeated measures ANOVA were performed to determine relationships between training variables and injury incidence over time.

To predict relationships between training load variables and injury risk over time, Generalised Additive Models were fitted in R (version 3.6.1, R Core Team, Austria (2019), <https://www.R-project.org/>) using the *mgcv*<sup>[1]</sup> package (version 1.8-31 for Generalized Additive Models).<sup>[13]</sup> Specifically, the relationship between training load variables (training distance, frequency, and the ACWR) and injury (dichotomised to yes/no) were modelled using Generalised Additive Models. Separate models were created for each training load variable (distance, frequency, and ACWR). All models specified a binomial distribution with a p-value <0.05 used to evaluate the significance of the effect predicted by the models. Similarly, Generalised Additive Models were fit to predict the relative risk for injury with training load variables; these required models to be calibrated as a Poisson distribution.

Relative risk (RR) with 95% CI was modelled in the Generalised Additive Models for training distance, frequency, and the ACWR. A RR of greater than 1 implied an increased risk for injury, while a RR of less than 1 indicated a decreased injury risk.

### Ethical considerations

The Faculty of Health Sciences Human Research Ethics Committee at the University of Cape Town (HREC/REF: 356/2018) gave ethical approval to conduct the study, and participants gave informed consent before their inclusion.

## Results

### Participants' characteristics

One hundred and fifty-six participants were recruited for this study. Thereafter, 50 participants were excluded based on the exclusion criteria (incomplete data), resulting in a sample of 106 participants. Descriptive characteristics are presented in Table 1.

A significant difference was found between the injured (n=43) and uninjured (n=63) groups in the number of Comrades

Ultramarathons completed ( $U_{(n=42, n=64)}=991.5$ ;  $z=-2.2$ ;  $p=0.03$ ) with the uninjured group having completed more Comrades ultramarathons as shown in Table 1.

**Injury profiles**

The prevalence of previous RRIs before the study period was 74% (n=78). In the injured group (n=43), 84% had a history of RRIs, compared to 67% in the uninjured group (n=63). A significant difference between the injured and uninjured groups ( $X^2=5.265$ ,  $p=0.02$ ) was found. Participants who reported a previous RRI were more likely to sustain an injury during the study compared to participants who had no history of previous injury.

Table 2 presents the total number of injuries (n=58) sustained during this study. Based on the IOC definitions, the injuries are classified as index, subsequent, exacerbation, or recurrent.<sup>[6]</sup> Exacerbations were not included in the total count as the definition implies that the injury occurred before the start of the study and had not fully healed yet.

**Weekly injury prevalence**

The highest weekly injury prevalence (13.2%) occurred during week 13, which was the week following the Comrades ultramarathon. The second highest injury prevalence (10.4%) was recorded in the first week of the study which was 12 weeks before the race. The average weekly prevalence of RRIs was 5% (mean  $4.9 \pm 3.4\%$ ).

**Anatomical region and types of injuries**

Muscles (47%) followed by injuries to tendons (24%) with an incidence of 4 and 2.5 injuries/1000 hours of running participation respectively were the most common structures injured. The knee (26%) followed by the hip (19%), and the foot (14%) were the most common anatomical structures reported as shown in Table 3.

**Training load profiles in relation to running-related injuries**

A significant difference in average weekly training frequency (sessions.wk<sup>-1</sup>) ( $t=3.27$ ,  $df=1467$ ,  $p=0.001$ ) and distance (km.wk<sup>-1</sup>) ( $t=2.71$ ,  $df=1468$ ,  $p=0.007$ ) was found between the two groups as shown in Table 4. The injured group had a lower average frequency and distance (km.wk<sup>-1</sup>) of running per week.

**Training distance, duration and frequency**

The weekly training distance of the injured group was consistently lower than the uninjured group ( $F_{(9, 926)}=12.35$ ;

$p=0.000004$ ) in the first 11 weeks before the Comrades ultramarathon (Figure 1).

As shown in Figure 1, the injured group had a lower average weekly training duration than the uninjured group ( $F_{(7, 101)}=7$ ;  $p=0.004$ ) in the first 11 weeks before the Comrades ultramarathon. The injured group consistently trained less often per week than the uninjured group ( $F_{(13, 1326)}=88.1$ ;  $p=0.00001$ ) over the 14-week period presented in Figure 1.

**Acute: chronic workload ratio**

Descriptive statistics for weekly average ACWR showed no significant difference between the injured and uninjured groups (Supplementary data).

**Predicting the relationship between training load parameters and RRI risk**

*a) The relationship between the ACWR and RRI risk*

A Generalised Additive Model was used to model the interaction between the ACWR and injury risk. The ACWR had no significant effect on injury risk ( $R^2=0.074$ ,  $df=1.00$ ,  $p=0.3$ ), although the Generalised Additive Model was able to model the non-linear u-shaped relationship with minimal increases in RR as shown in Figure 2.

Data with ACWR >2.5 was limited in our study, resulting in greater uncertainty of the model over the ACWR range of 2.5. A relationship between lower and higher ACWRs and increased risk is shown by the Generalised Additive Model in Figure 2a. However, the magnitude of increased injury risk in this model was minimal with a maximum increase in RR of 0.7 which was not statistically significant.

*b) The relationship between weekly training distance and RRI risk*

The Generalised Additive Model depicting the relationship

**Table 1. Descriptive characteristics and endurance running participation history of participants grouped in total, injured and uninjured groups**

	Total (n=106)	Injured (n=43)	Uninjured (n=63)	t-value	p-value
Age (years)	40.5±7.9	40±6.5	40.8±8.8	-0.54	0.59
Height (cm)	172.2±11.1	170.4±11.3	173.4±10.9	-1.36	0.18
Mass (kg)	71.3±13.0	71.5±11.4	71.1±14.0	0.16	0.87
Participation in endurance running (years)	9.7±7.5	9.7±6.8	9.7±7.9	0.001	1.00
Ultramarathons completed (n)	8±12	7±10	9±13	1085 (U-value)	0.09
Comrades ultramarathons completed (n)	3±4	4±2	5±4	992 (U-value)	0.03*

\* $p<0.05$ . Data are expressed as mean ± standard deviation (SD).

**Table 2. Summary of injury count (n), prevalence (%), incidence (/1000h) and severity in time loss**

	Injuries n	Prevalence (%)	Incidence (/1000 hours)	Time-loss (days) Median (range)
Index injury	43	40.7%	6.3	4 (0-25)
Subsequent injury	12	9.4%	1.8	3 (0-9)
Exacerbation injury	9	5.7%	1.3	3 (0-9)
Recurrent injury	3	2.8%	0.4	2 (0-4)

between the training load variable of training distance ( $\text{km.wk}^{-1}$ ) and injury risk indicated a more linear relationship. A significant inverse linear relationship was found ( $R^2=2.311$ ,  $df=1165$ ,  $p=0.02$ ).

A stable injury risk was shown when the weekly training distance was just over  $100 \text{ km.wk}^{-1}$  ( $RR_{\text{Distance}=107}=1$ ,  $p<0.05$ ). The RR was the highest at the lowest training distance ( $RR_{\text{Distance}=0}=2.3$ ,  $p<0.05$ ).

A decrease in relative risk was then observed as the training distance increased, reaching the lowest relative risk at a training distance of just over  $150 \text{ km.wk}^{-1}$  ( $RR_{\text{Distance}=160}=0.8$ ,  $p<0.05$ ) as presented in Figure 2b.

### c) The relationship between weekly training frequency and RRI risk

The Generalised Additive Model predicting the relationship between injury and frequency of running training per week showed a significant effect ( $R^2=3.293$ ,  $df=2.62$ ,  $p=0.02$ ).

The RR of injury was shown to be the highest with a training frequency of  $<2$  sessions per week with a relative risk of injury just under 3 ( $RR_{\text{Frequency}=0}=3$ ,  $p<0.05$ ;  $RR_{\text{Frequency}=2}=3.3$ ,  $p<0.05$ ). With a training frequency of  $>2$  sessions. $\text{wk}^{-1}$  the RR for injury declines, until a stable injury risk was shown with a weekly training frequency of five sessions per week ( $RR_{\text{Frequency}=5}=1$ ,  $p<0.05$ ). Training more than five sessions per week was shown to be protective against injury. The lowest RR is shown at a training frequency of eight sessions. $\text{wk}^{-1}$  ( $RR_{\text{Frequency}=8}=0.25$ ,  $p<0.05$ ) as shown in Figure 2c

## Discussion

In this study, we aimed to investigate the incidence and nature of RRIs and their association with training loads in runners 12 weeks before and two weeks after the 2018 Comrades ultramarathon. A wide range in prevalence and incidence of RRIs across studies exists which is likely due to differences in injury definitions, sample size, and the specific population studied.<sup>[2-5,7,17,18]</sup> We observed a higher injury rate with an overall incidence of 8 injuries per 1000 hours compared to 5 injuries per 1000 hours of running training reported by Craddock et al.<sup>[12]</sup>, and lower than the average injury rates reported by Hoffman and Fogard<sup>[17]</sup> and Hoffman and Krishnan<sup>[5]</sup>. The higher injury rate reported in the present study may be explained by the longer event distance the participants trained for compared to other studies.<sup>[12,19]</sup>

Two retrospective studies on ultramarathon runners by Hoffman and Fogard<sup>[17]</sup> and Hoffman and Krishnan<sup>[5]</sup> reported a higher RRI proportion prevalence of 65% and 52%, respectively, compared to 40% in this study. These studies had a larger sample size, which may have resulted in greater statistical power. The authors also assessed runners over 12

**Table 3. Injury distribution, incidence (n/1000 hours) and time-loss (days) of all recorded injuries**

Injury description – Anatomical region / type of structure	Injuries n	Incidence Injuries/1000 hours	Time-loss Days (95% CI)
<b>Total injuries by structure type</b>			
Muscle	27	3.7	2.9 (0-9)
Tendon	14	2.5	4.5 (0-25)
Ligament	3	0.7	3.5 (2-5)
Joint	2	0.3	3 (0-6)
Bone	2	0.4	1 (0-2)
Nerve	3	0.4	4 (3-5)
Other	8	1.5	3.3 (0-9)
<b>Injuries by region and structure type</b>			
<b>Back</b>	<b>2</b>	<b>0.3</b>	<b>4 (2-6)</b>
Muscle	1	0.1	2 (2-2)
Joint	1	0.1	6 (6-6)
<b>Hip</b>	<b>11</b>	<b>1.6</b>	<b>2.6 (0-6)</b>
Muscle	4	0.6	1.3 (0-4)
Tendon	4	0.6	4.3 (0-6)
Bone	1	0.1	2 (2-2)
Nerve	1	0.1	5 (5-5)
Do not know	1	0.1	2 (2-2)
<b>Quadriceps (front of thigh)</b>	<b>6</b>	<b>0.9</b>	<b>4.9 (2-13)</b>
Muscle	6	0.9	4.8 (2-11)
Tendon	1	0.1	2 (2-2)
<b>Hamstrings (back of thigh)</b>	<b>3</b>	<b>0.4</b>	<b>3 (2-4)</b>
Muscle	3	0.4	3 (2-4)
<b>Groin</b>	<b>2</b>	<b>0.3</b>	<b>7.5 (6-9)</b>
Tendon	2	0.3	7.5 (6-9)
<b>Knee</b>	<b>15</b>	<b>2.2</b>	<b>3 (0-9)</b>
Muscle	3	0.4	3 (0-5)
Joint	1	0.1	0 (0-0)
Tendon	3	0.4	1.5 (0-3)
Ligament	2	0.3	3 (2-4)
Nerve	2	0.3	3.5 (3-4)
ITB	3	0.4	15 (9-20)
Do not Know	1	0.1	0 (0-0)
<b>Calf</b>	<b>7</b>	<b>1</b>	<b>2.8 (0-4)</b>
Muscle	6	0.7	2.5 (0-4)
Tendon	1	0.1	4 (4-4)
<b>Shin</b>	<b>1</b>	<b>0.1</b>	<b>0 (0-0)</b>
Muscle	1	0.1	0 (0-0)
<b>Ankle</b>	<b>3</b>	<b>0.4</b>	<b>5 (4-6)</b>
Muscle	1	0.1	4 (4-4)
Tendon	1	0.1	6 (6-6)
Ligament	1	0.1	5 (5-5)

months, while our study evaluated only for injuries sustained during 14 weeks.

The most prevalent type of structures injured were muscles (3.7/1000h) followed by injuries to tendons (2.5/1000h). Muscle injuries had a higher incidence in the quadriceps, calf, and hip region, while tendon injuries had a higher incidence in the hip and knee region, similar to related studies.<sup>[5,12,17]</sup>

The knee (26%) was reported as the most injured anatomical area, similar to Craddock et al.<sup>[12]</sup> with 19%, and in line with numerous studies ranging between 7% and 50%.<sup>[4,18,19]</sup> Due to the greater burden that injuries to the knee and foot have on ultramarathon runners, as reported by this study, further research should focus on possible aetiological factors regarding the development of knee and foot injuries in ultramarathon runners.

The severity of injuries differed according to the type of structure injured, with tendon injuries resulting in the most time loss. Compared to related studies, we found the median injury severity according to time-loss in days to be four days, aligned with Hoffman and Krishnan<sup>[5]</sup> who reported a median of five days' time-loss.

The relationship between training load and injury risk has long been proposed to be non-linear.<sup>[20]</sup> Recently, there has been a call for better handling non-linearity in sports medicine research.<sup>[14,15]</sup> Statistical methods commonly used in training load research have been criticised for not accounting for repeated measures.<sup>[13-15]</sup> This has been addressed by using General Linear Models in certain cases and discretising the training load parameters for analysis.<sup>[13]</sup>

Generalised Linear Models are not ideally suited as they assume that the relationship between the training load variables (external training load and ACWR) and the outcome variable (injury status) is constant over the exposure range.<sup>[13]</sup> It has also been argued that data discretisation is inappropriate for continuous data and can result in inflated discovery rates or negative results.<sup>[15]</sup> It has been recommended that the methods used in future research should be able to explore and model non-linearity to discover the relationship and should not constrain it.<sup>[14]</sup> Further, it has been proposed that modelling training load and injury risk in GAMs might be better suited to model the flexible relationship across the diverse exposure range of training load and injury risk.<sup>[10,13]</sup>

The present study modelled training load variables using traditional statistical modelling methods and Generalised Additive Models. The modelled relationship between injury risk and weekly training load confirms our finding that a

**Table 3 Continued. Injury distribution, incidence (n/1000 hours) and time-loss (days) of all recorded injuries**

Injury description – Anatomical region / type of structure	Injuries n	Incidence Injuries/1000 hours	Time-loss Days (95% CI)
<b>Injuries by region and structure type</b>			
Foot	8	1.2	5.7 (0-25)
Muscle	2	0.3	4.5 (4-5)
Tendon	2	0.3	9 (3-15)
Bone	1	0.1	0 (0-0)
Skin	1	0.1	5 (5-5)
Plantar Fasciitis	1	0.1	2 (2-2)
Do not know	1	0.1	3 (3-3)

lower weekly training mileage is associated with an increased injury risk, which is supported in related literature.<sup>[12,18]</sup> A significant inverse linear relationship between a lower training distance and greater injury risk was reported using a Generalised Additive Model. The relative risk for injury gradually increased as the weekly mileage decreased, reaching a two-fold increase in injury risk in participants who trained less than 25km per week compared to participants who trained 100km per week. Craddock et al.<sup>[12]</sup> reported that a weekly training distance of less than 30km per week was a risk factor for injury in ultramarathon and marathon participants, respectively. It is reasonable to argue that participants might have a more cautious approach to training following an injury. This could have led to a reduction in the weekly training load as a result of the injury.

A stable injury risk (neither an increase nor decrease in risk) was reported when the weekly training distance was slightly over 100km per week. A maximal reduction in injury risk was found when the training distance was approximately 150km per week. Research has suggested that an upper limit of training volume exists, at which point an increase in injury risk transpires.<sup>[3,8]</sup> We found that such an upper limit may be much higher in the population of ultramarathon runners, as a reduced injury risk with a weekly training distance as high as 180km per week was predicted using the Generalised Additive Model.

Research on training load injury has been criticised for numerous conceptual and methodological pitfalls.<sup>[21]</sup> More specifically, the overreliance on the training load metric of ACWR in related literature has been criticised.<sup>[21]</sup> The lack of a conceptual framework to support the theoretical mechanisms that might explain the influence of ACWR on health and performance outcomes has been noted.

**Table 4. Training load variables (average duration, average distance, average frequency, average (ACWR) during the 14-week study period**

	Total (n=106)	Injured (n=42)	Uninjured (n=64)	t-value	df	p-value	Cohen's d effect size
Average Frequency (sessions.wk <sup>-1</sup> )	3.0±2.0	2.9±1.9	3.3±1.9	3.27	1467	0.001*	0.17
Average Distance (km.wk <sup>-1</sup> )	54±37	51±39	56±36	2.71	1468	0.007*	0.14
Average Duration (min.wk <sup>-1</sup> )	335±239	321±264	345±219	1.91	1467	0.06	0.10
Average ACWR (AU)	0.9±0.6	0.9±0.6	0.9±0.6	-0.72	1163	0.53	-0.04

\*p<0.05. Data presented in total, injured and uninjured groups and expressed as mean±standard deviation (SD); ACWR, acute-to-chronic workload ratio

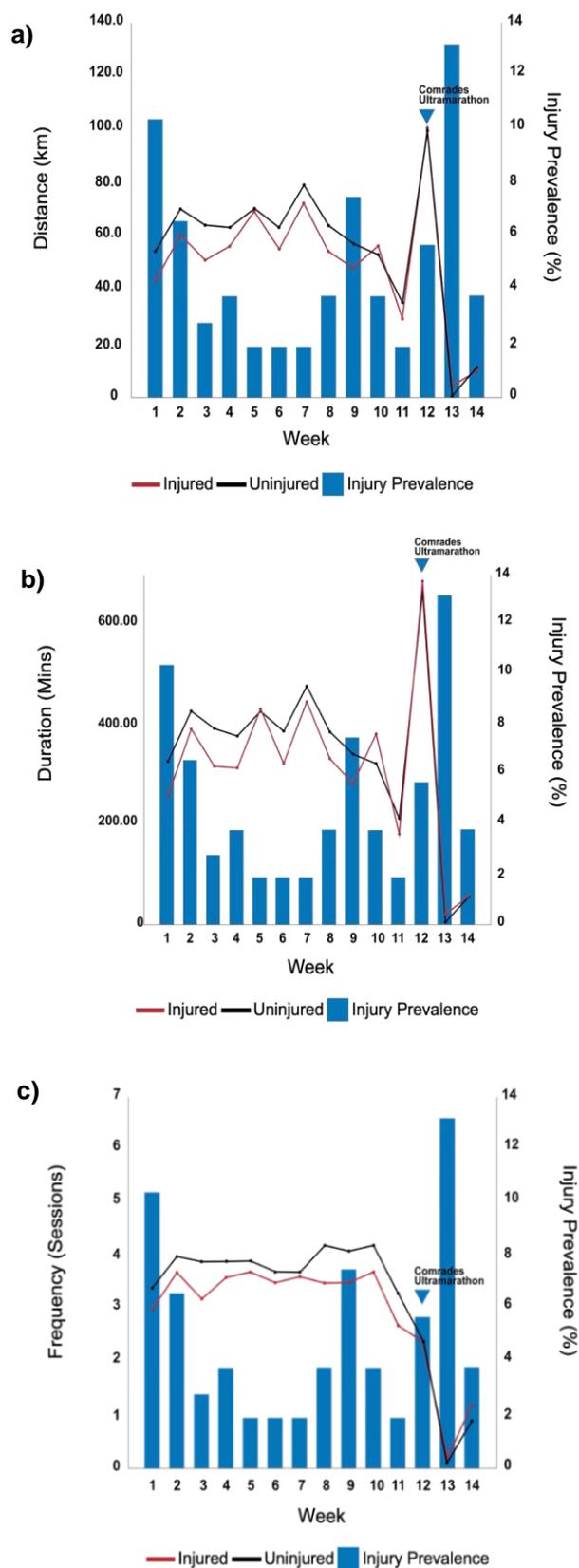


Fig. 1. Average weekly training (a) distance, (b) duration, (c) frequency between the injured and uninjured group over the 14-week study period in kilometres. The average weekly prevalence of injuries (%) is presented.

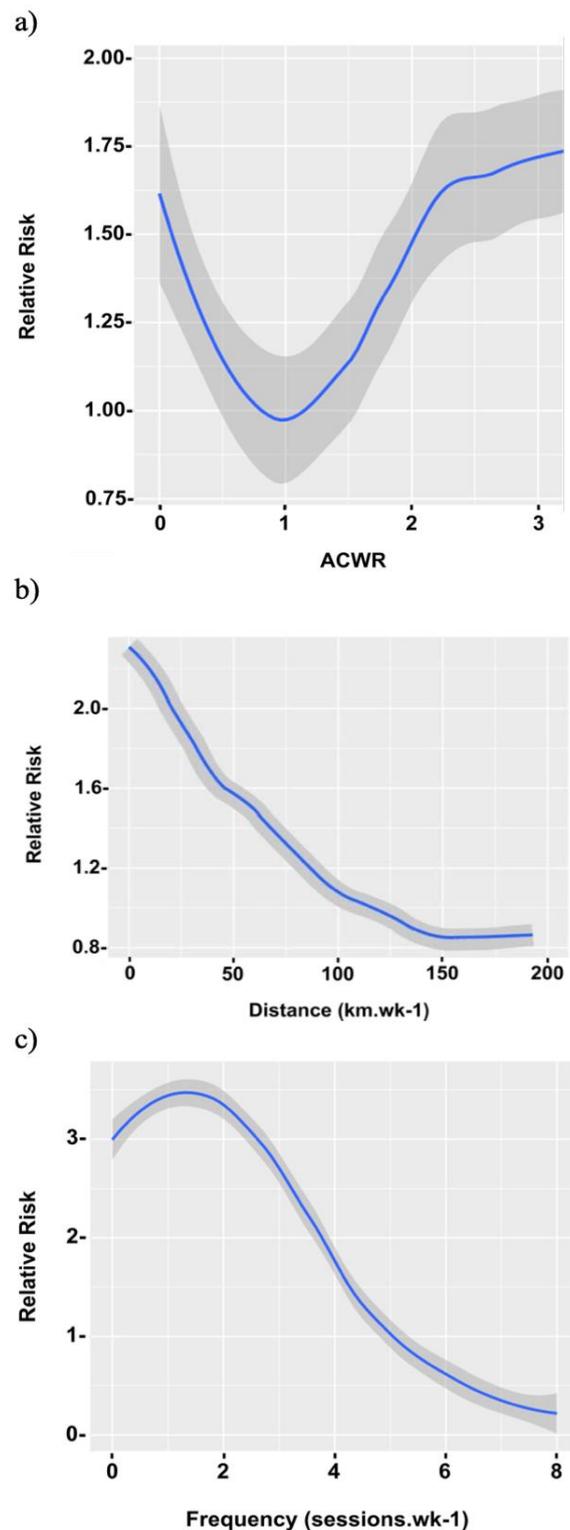


Fig. 2. Generalised Additive Models illustrating the relationship between a) the 10-week acute-to-chronic workload ratio (ACWR) and injury risk in ultramarathon runners, b) the 14-week training load variable of distance (km.wk<sup>-1</sup>) and injury in ultramarathon runners, and c) the 14-week training load variable of frequency of training (sessions.wk<sup>-1</sup>) and injury in ultramarathon runners. The blue line represents the model with 95% CI in the shaded area. The relative risk of injury is shown on the y-axis.

Using the Generalised Additive Model, our study modelled the relationship between injury risk and ACWR and illustrated a heterogeneous relationship. The effect of weekly ACWR on injury risk was not significant in our study population. To our knowledge, the present study is the first to use Generalised Additive Models to predict the relationship between injury risk and the ACWR in ultramarathon runners. We cannot compare the present findings with the current literature. Only two other studies used Generalised Additive Models to model ACWR data with injury risk.<sup>[10,13]</sup> In these studies, injury risk was investigated in children based on changes in physical activity over time.<sup>[10,13]</sup> The ACWR with injury risk modelled as a Generalised Additive Model was a better predictor of injury risk when compared to the General Linear Model.<sup>[10,13]</sup>

The Generalised Additive Models used in the present study illustrate a U-shaped relationship between injury risk and ACWR. The Generalised Additive Model predicted a stable injury risk with an ACWR of between  $\sim 0.8$  and  $1.1$ . Both an ACWR of  $<0.8$  and  $>1.1$  resulted in an increased RR for injury, with the highest RR at an ACWR=0 and ACWR=2.2, similar to findings by Wang et al.<sup>[13]</sup> Our model predicted that the increase in injury risk was minimal with the increase in ACWR, much lower than what is predicted by the IOC's model.<sup>[20]</sup>

Our findings using a Generalised Additive Model closely resembled the proposed model by the IOC on the relationship between changes in training load and injury risk in athletes.<sup>[20]</sup> Due to the non-linear nature of the proposed relationship between ACWR and injury risk, a Generalised Additive Model may provide better insight into the predictive nature of ACWR on injury risk in athletes in future studies.

Regarding the practical implications of these findings, there seems to be a clear indication that insufficient training/low training load during the final 12 weeks before an ultramarathon increases the risk of sustaining an injury. While other possible confounding factors, like speed, were not considered (many speed sessions may be performed at higher velocities than longer duration runs), the findings do support previous hypotheses that low training loads may not sufficiently prepare runners for the demands of the long duration and distance in ultramarathons.

### Limitations

While a modest sample size ( $n=106$ ) was used in this study, similar studies included a participant range between 32 and 873.<sup>[9]</sup> Self-reporting of the type of injury may be unreliable as it depends on participants' subjective understanding of their injury. At the same time, the reporting of anatomical areas seems to be more reliable.<sup>[5]</sup>

The study's retrospective design may have led to recall bias of when an injury occurred. Individuals were recruited within five months after the 2018 Comrades ultramarathon to minimise recall bias. Data obtained retrospectively within 12 months has a minimal recall bias of approximately 12%.<sup>[22]</sup> In addition, the researchers facilitated the completion of the questionnaires either in person or online to clarify any questions that arose which would have prevented the

participants from completing the questionnaires due to confusion. Using the objective data from the GPS devices was an attempt to reduce the risk of recall bias related to the distance, duration and frequency. Still, it resulted in other potential biases due to the economic burden of such devices in a country like South Africa with significant socioeconomic disparities.

Our study is unique because it obtained training data stored on GPS watches. All runners were required to have a GPS watch to participate in this study. However, we recognise that GPS watches can be costly and may not be accessible to many runners in South Africa, potentially skewing the eligibility of numerous runners who would otherwise have been suitable for the study and contributing to a more diverse population. Future studies should include GPS watches as a standard line item for research funding to ensure research is accessible by a representative sample of runners.

### Conclusion

The overall injury incidence reported in this study was comparable to that of the existing literature. The proportional prevalence of running-related injuries was lower than previously reported. Lower training loads were associated with a higher risk for RRI in ultramarathon runners. The frequency of running training per week substantially influenced injury risk. Sudden changes in training load, as calculated by the ACWR, appeared to have a minimal effect on injury risk in the ultramarathon runners' population.

The non-linear relationship between several training load variables and injury risk can successfully be modelled with Generalised Additive Models. Future research is required to establish whether the predicted parameters of training load variables may assist in preventing injury in ultramarathon runners in well-designed randomised controlled trials.

**Conflict of interest and source of funding:** The authors declare that they have no personal or financial relationships that may have influenced in the writing of this article. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors

**Acknowledgements:** We would like to thank Curt Barnes for his contributions to editing the final document.

**Author contributions:** TB, PD, and KB made substantial contributions to the development of the study, data collection, analysis, and write-up of this manuscript, as well as the criteria for authorship as outlined by the journal. All authors approved the final version.

### References

- Hoffman MD. Injuries and health considerations in ultramarathon runners. *Phys Med Rehabil Clin N Am.* 2016;27(1):203-216. [<http://dx.doi.org/10.1016/j.pmr.2015.08.004>] [PMID: 26616184]
- Kakouris N, Yener N, Fong DTP. A systematic review of running-related musculoskeletal injuries in runners. *J Sport Health Sci.* 2021;10:513-522. [<http://dx.doi.org/10.1016/j.jshs.2021.04.001>] [PMID: 33862272]

3. Saragiotto BT, Yamato TP, Hespanhol Junior LC, Rainbow MJ, Davis IS, Lopes AD. What are the main risk factors for running-related injuries? *Sports Med.* 2014;44:1153-63. [<http://dx.doi.org/10.1007/s40279-014-0194-6>][PMID: 24809248]
4. Videbæk S, Bueno AM, Nielsen RO, Rasmussen S. Incidence of running-related injuries per 1000 h of running in different types of runners: a systematic review and meta-analysis. *Sports Med.* 2015;45(7):1017-26. [<http://dx.doi.org/10.1007/s40279-015-0333-8>] [PMID: 25951917]
5. Hoffman MD, Krishnan E. Health and exercise-related medical issues among 1,212 ultramarathon runners: Baseline findings from the Ultrarunners Longitudinal TRacking (ULTRA) Study. *PLoS one.* 2014;9(1):e83867. [<http://dx.doi.org/10.1371/journal.pone.0083867>][PMID: 24416176]
6. Bahr R, Clarsen B, Derman W, Dvorak J, Emery C, Finch C, et al. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sport (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). *Br J Sports Med.* 2020. [<http://dx.doi.org/10.1177/2325967120902908>][PMID: 32071062]
7. Hulme A, Nielsen RO, Timpka T, Verhagen E, Finch C. Risk and protective factors for middle- and long-distance running-related injury. *Sports Med.* 2017;47(5):869-86. [<http://dx.doi.org/10.1007/s40279-016-0636-4>][PMID: 27785775]
8. Nielsen RØ, Parner ET, Nohr EA, Sørensen H, Lind M, Rasmussen S. 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(10):739-47. [<http://dx.doi.org/10.2519/jospt.2014.5164>]
9. Johnston R, Cahalan R, O'Keefe M, O'Sullivan K, Comyns T. The associations between training load and baseline characteristics on musculoskeletal injury and pain in endurance sport populations: a systematic review. *J Sci Med Sport.* 2018;21:910-8. [<http://dx.doi.org/10.1016/j.jsams.2018.03.001>][PMID: 29559317]
10. Wang C, Stokes T, Vargas JT, Steele R, Wedderkopp N, Shrier I. Predicting injury risk over changes in physical activity in children using the acute:chronic workload ratio. *Am J Epidemiol.* 2022;191(4):665-73. [<http://dx.doi.org/10.1093/aje/kwab280>] [PMID: 34849538]
11. Gabbett TJ. The training-injury prevention paradox: should athletes be training smarter or harder? *Br J Sports Med.* 2016;50:273-80. [<http://dx.doi.org/10.1136/bjsports-2015-095788>] [PMID: 26758673]
12. Craddock NL, Buchholtz K, Burgess TL. Does a greater training load increase the risk of injury and illness in ultramarathon runners? A prospective, descriptive, longitudinal design. *S Afr J Sports Med.* 2020;32(1):1-6. [<http://dx.doi.org/10.17159/2078-516X/2020/v32i1a8559>] [PMID: 36818970] [PMCID: PMC9924505]
13. Wang C, Stokes T, Vargas JT, Steele R, Wedderkopp N, Shrier I. Injury risk increases minimally over a large range of changes in children. *arXiv. preprint arXiv:2010.02952.* 2020 Oct 6. [<http://doi.org/10.48550/arXiv.2010.02952>]
14. Bache-Mathiesen LK, Andersen TE, Dalen-Lorentsen T, Clarsen B, Fagerland MW. Not straightforward: modelling non-linearity in training load and injury research. *BMJ Open Sport Exerc Med.* 2021;7(3): e001119. [<http://dx.doi.org/10.1136/bmjsem-2021-001119>][PMID: 34422292]
15. Carey DL, Crossley KM, Whiteley R, Mosler A, Ong KL, Crow J, et al. Modeling training loads and injuries: The dangers of discretization. *Med Sci Sports Exerc.* 2018;50(11):2267-76. [<http://dx.doi.org/10.1249/MSS.0000000000001685>] [PMID: 29933352]
16. Mundo AI, Tipton JR, Muldoon TJ. Using generalized additive models to analyze biomedical non-linear longitudinal data. *bioRxiv.* Jun 11:2021-06. [<http://dx.doi.org/10.1101/2021.06.10.447970>]
17. Hoffman MD, Fogard K. Factors related to successful completion of a 161-km ultramarathon. *Int J Sports Physiol Perform.* 2011;6(1):25-37. [<http://dx.doi.org/10.1123/ijspp.6.1.25>] [PMID: 21487147]
18. Van Gent RN, Siem D, van Middelkoop M, Van Os AG, Bierma-Zeinstra SMA, Koes BW. Incidence and determinants of lower extremity running injuries in long distance runners: A systematic review. *Br J Sports Med.* 2007;41(8):469-80. [<http://dx.doi.org/10.1136/bjsem.2006.033548>][PMID: 17473005]
19. Kluitenberg B, van Middelkoop M, Diercks R, van der Worp H. What are the differences in injury proportions between different populations of runners? A systematic review and meta-analysis. *Sports Med.* 2015;45(8):1143-61. [<http://dx.doi.org/10.1007/s40279-015-0331-x>] [PMID: 25851584]
20. Soligard T, Schweltnus M, Alonso JM, Bahr R, Clarsen B, Dijkstra HP 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(17):1030-41. [<http://dx.doi.org/10.1136/bjsports-2016-096581>] [PMID: 27535989]
21. Impellizzeri FM, McCall A, Ward P, Bornn L, Coutts AJ. Training load and its role in injury prevention, part 2: conceptual and methodologic pitfalls. *J Athl Train.* 2020;55(9):893-901. [<http://dx.doi.org/10.4085/1062-6050-501-19>] [PMID: 32991699]
22. Lopes AD, Hespanhol LC, Yeung SS, Pena Costa LO. What are the main running-related musculoskeletal injuries? *Sports Med.* 2012;42(10): 891-905. [<http://dx.doi.org/10.2165/11631170-000000000-00000>][PMID: 22827721]