Original research

A step towards understanding the mechanisms of running-related injuries

Laurent Malisoux a,*, Rasmus Oestergaard Nielsen b, c, Axel Urhausen a, d, Daniel Theisen a

a Sports Medicine Research Laboratory, Public Research Centre for Health, Luxembourg
b Department of Public Health, Section of Sport Science, Aarhus University, Denmark
c Orthopedic Surgery Research Unit, Science and Innovation Center, Aalborg University Hospital, Denmark
d Sports Clinic, Centre Hospitalier de Luxembourg, Luxembourg

A R T I C L E   I N F O

Article history:
Received 5 June 2014
Received in revised form 2 July 2014
Accepted 25 July 2014
Available online 12 August 2014

Keywords:
Sports injury prevention
Training load monitoring
Effect-measure modification
Injury mechanism

A B S T R A C T

Objectives: To investigate the association between training-related characteristics and running-related injury using a new conceptual model for running-related injury generation, focusing on the synergy between training load and previous injuries, short-term running experience or body mass index (> or <25 kg m−2).

Design: Prospective cohort study with a 9-month follow-up.

Methods: The data of two previous studies using the same methodology were revisited. Recreational runners (n = 517) reported information about running training characteristics (weekly distance, frequency, speed), other sport participation and injuries on a dedicated internet platform. Weekly volume (dichotomized into <2 h and ≥2 h) and session frequency (dichotomized into <2 and ≥2) were the main exposures because they were considered necessary causes for running-related injury. Non-training-related characteristics were included in Cox regression analyses as effect-measure modifiers. Hazard ratio was the measure of association. The size of effect-measure modification was calculated as the relative excess risk due to interaction.

Results: One hundred sixty-seven runners reported a running-related injury. Crude analyses revealed that weekly volume <2 h (hazard ratio = 3.29; 95% confidence intervals = 2.27; 4.79) and weekly session frequency <2 (hazard ratio = 2.41; 95% confidence intervals = 1.71; 3.42) were associated with increased injury rate. Previous injury was identified as an effect-measure modifier on weekly volume (relative excess risk due to interaction = 4.69; 95% confidence intervals = 1.42; 7.95; p = 0.005) and session frequency (relative excess risk due to interaction = 2.44; 95% confidence intervals = 0.48; 4.39; p = 0.015). A negative synergy was found between body mass index and weekly volume (relative excess risk due to interaction = −2.88; 95% confidence intervals = −5.10; −0.66; p = 0.018).

Conclusions: The effect of a runner’s training load on running-related injury is influenced by body mass index and previous injury. These results show the importance to distinguish between confounding and effect-measure modification in running-related injury research.

© 2014 Sports Medicine Australia. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

Risk factors for running-related injury (RRI) in runners have been widely investigated.1–3 Such studies are extremely valuable to identify populations at risk. In spite of past research efforts, only few consistent risk factors have been revealed in the literature, probably due to different study designs and analytical approaches used.5 Moreover, the sole identification of risk factors is insufficient to elucidate the mechanisms involved in RRI generation,6,7 a prerequisite for successful injury prevention measures.1,8

To date, evidence on RRI aetiology is virtually non-existent. One of the main reasons regularly highlighted is the absence of large-scale prospective cohort studies.6,8 In addition, the conceptual and statistical approach used for data-analysis has been given insufficient attention. The classical way used by many authors is to run regression analyses,10,11,13 where all variables thought to be related to injury are first tested separately for their association with RRI. Next, those below a certain p-value are included in a final adjusted model. This approach implies that each included variable is a confounder for the outcome and is directly associated with it.
Personal characteristics such as age, body mass index (BMI), previous injury, preferred running surface or use of different pairs of running shoes have previously been suggested to be related to RRI. \(^1\) Another study concluded that atypical foot pronation and inadequate hip muscle stabilization were suspected mechanisms involved in the cause of overuse running injuries. \(^3\) However, strictly speaking, none of these factors in themselves are sufficient causes for injury. Runners do not sustain an RRI only because they are overweight, older, or have had a previous injury. \(^4, 14\) RRIs can only occur when people practice running. \(^16\) This means that running practice is a necessary cause for RRI and, in fact, the only necessary cause. Therefore, when studying causal mechanisms, training-related characteristics should be considered as primary exposures of interest in RRI research. Unfortunately, there is so far only limited evidence about the association between training-related characteristics and RRI. \(^1, 18\) Previously, experts have argued that half of all RRIs are related to training errors and could be preventable. \(^3, 10\) However, a systematic review failed to identify which of these training errors are related to RRI. \(^9\)

The identification of training errors represents a particularly interesting line of attack regarding injury prevention. On the one hand, running-related characteristics (e.g., training volume and frequency) are necessary factors for injury development, \(^4\) on the other hand a runner’s training regime is easily modified. \(^17\) However, the mere detection of risk factors without understanding the underlying mechanism is insufficient to optimize prevention measures. \(^7\) Although a multifactorial model for sports injury aetiology was suggested already 20 years ago, \(^14\) no study to date has investigated if and how injury predictors work in synergism. To fill this knowledge gap, we suggest here a conceptual model of RRI generation in which the primary exposures of interest are training-related characteristics. Non-training-related characteristics are considered as potential effect-measure modifiers (EMM) because the effect of training-related exposure is different across strata of non-training-related factors. Therefore, the aim of this study was to investigate the influence of training volume (hours per week) and training frequency on RRI, and especially to analyse in how far other personal characteristics affect these relationships. To achieve this goal, we combined and re-analysed the data of over 500 recreational runners collected in the framework of 2 previous studies. \(^11, 12\)

2. Methods

A prospective observational study \(^11\) and a randomized control trial \(^2\) were initiated in parallel in 2012, using the same methodology. All participants (above 18 years) signed an informed consent and were free to follow their own training programme. The main study requirements in both studies were: (1) to train on average at least once a week, (2) to report training data related to running and all other sporting activities (per training session) at least once a week, (3) to systematically report any injury and illness sustained during the 9-month follow-up period, (4) to have no contraindication to running training (e.g., injury) at the time of initial inclusion, and (5) to have no degenerative conditions and no history of surgery to the lower limbs or the back region within the previous 12 months. The study protocols and online procedures had previously been approved by the National Ethics Committee for Research (Refs. 201111/10 and 201201/02).

A total of 754 participants initially created their account on the Training and Injury Prevention Platform for Sport (TIPPS) website during the recruitment phase of the randomized controlled trial \((n = 299)\) and the observational study \((n = 455)\). The demographic data gathered were: age, sex, weight, height, regular running practice over the previous 12 months (number of months with at least one session a week), running experience (years of previous regular practice) and previous (12 months) injury to the lower back or lower limbs preventing the participant from normal running activity.

The injury definition was a modified version of the one used by Buist et al. \(^2\): any physical pain located at the lower limbs or lower back region, sustained during or as a result of running practice and impeding planned running activity for at least 1 day (time-loss definition). Participants were instructed to report all adverse events including injuries preventing them from normal running activity via a dedicated questionnaire on their TIPPS account. In the present study, overuse and traumatic non-contact injuries were included in the analyses, whatever the mode of onset (sudden or gradual). \(^19\) RRIs were classified according to consensus guidelines on sports injury surveillance studies. \(^19, 20\)

During follow-up, participants were instructed to upload all running sessions and other sporting activities undertaken onto their TIPPS account. \(^21\) Primary exposures were weekly running volume and weekly session frequency. Running practice characteristics were described as average values during the follow-up period. Dichotomization was done for weekly running volume \((<2\text{h} \text{and} ≥2\text{h week}^{-1})\) and weekly session frequency \((<2\text{ and} ≥2\text{ sessions week}^{-1})\), based on the respective median.

Individual e-mail reminders were sent to the participants who had not provided the system with any data for the previous week. Injury data was systematically checked by one of the investigators for completeness and coherence. Personal phone calls were made if the reported information on the injury form was found to be inconsistent. A participant was considered as dropping out of the study when no data was uploaded in the system for more than 2 weeks despite the automatic reminder sent by the system and a phone call from the research team.

Effect-measure modifiers were BMI, previous injury and short-term (12 previous months) regular running experience. BMI was dichotomized into <25 and ≥25 kg m\(^{-2}\). Runners were considered as regulars if they had practiced running on a weekly basis over the previous 12 months. Previous injury was defined as any RRI sustained over the previous 12 months.

Cox regression was used to compute the hazard rates in the exposure groups, using RRI as the primary outcome and hours spent running (time at risk, expressed in hours) as the time-scale. \(^22\) Date at inclusion and date at injury (if applicable) or at censoring were basic data used to calculate the time at risk. Participants were right-censored in case of severe disease, non-running-related injury causing a modification of the running plan or at the end of follow-up, whichever came first. The assumption of proportional hazards was evaluated by log-minus-log plots to validate the statistical model. In addition, the recommendation of using at least 10 injuries per predictor variable included in the Cox regression analysis was followed strictly. \(^23\)

As a preliminary phase, unadjusted Cox proportional hazard regressions were performed to present the crude estimates of training-related characteristics. To study whether the effects of the primary exposures on RRI were modified by previous injury, short-term running experience and BMI (cf. Fig. 1), the additional following steps were performed, according to the recommendations by Knol and VanderWeele. \(^24\) First, stratified analyses were performed separately for each of the two training characteristics (weekly volume and frequency) including either previous injury, short-term running experience or BMI as potential EMM (thus creating 4 strata for each analysis). Hazard ratios (HR) and their 95% confidence intervals (95%CI) were determined for each stratum with a single reference category (the stratum with the lowest injury rate). Secondly, HR and the corresponding 95%CI were computed within strata of previous injuries, short-term running experience and BMI. Finally, the size of the effect-measure modification was
calculated as the relative excess risk due to interaction (RERI), using the additive scale. Synergism between two exposures was concluded if 0 was not comprised in the 95%CI of the RERI. An RERI value above zero implies a positive synergism while a negative value implies a negative synergism.

Cut-off values for dichotomization were determined, amongst others, with the aim to get at least 15 participants with and without injuries within each of the strata. Significance was accepted for \( p < 0.05 \). In addition, estimated effect size and estimated precision (95% confidence limits) were used for proper interpretation of study results. All analyses were performed using SPSS V20.

3. Results

Of the 754 volunteers who initially registered to the prospective cohort study or the RCT, 237 of them were excluded from the analyses because they did not upload any sporting activity during the observation period, they reported <2 running sessions before the first RRI or censoring, or they did not provide all required information. Thus, a total of 517 recreational runners were eventually included in the analyses. Participants reported an average of 2.1 ± 1.1 running sessions per week, with a total volume of 2.3 ± 1.6 h week\(^{-1}\). Their mean running distance was 22.1 ± 16.2 km week\(^{-1}\) and the average running speed was 9.6 ± 1.6 km h\(^{-1}\). Personal and sport-related characteristics of the participants are presented in Table 1.

A non-contact RRI was sustained by 167 of the 517 participants (32.3%). For comparison purposes to previous studies, the overall incidence was 6.68 RRI/1000 h of running. Acute non-contact injuries (e.g. muscles tear) accounted for 13.8% (\( n = 23 \)) of the RRs, and 32.9% (\( n = 55 \)) of all injuries were recurrent. Most of the RRs affected muscles (44.9%) and tendons (41.3%), and the most often concerned anatomical locations were the lower leg (22.7%), the knee (22.2%) and the thigh (20.9%).

A crude analysis (unadjusted Cox regression model) of the association between the factors presented in Fig. 1 revealed that weekly volume <2 h (HR = 3.29; 95%CI = 2.27; 4.79) and session frequency <2 sessions per week (HR = 2.41; 95%CI = 1.71; 3.42), were associated with increased injury rate.

A stratified analysis according to previous injury is presented in Table 2. In both strata, the rate at which RRI occurred at any time was higher amongst the participants with a weekly volume <2 h and those who ran <2 sessions week\(^{-1}\). Moreover, previous injury was identified as an EMM, since the RERI on weekly volume (RERI = 4.69; 95%CI = 1.42; 7.95; \( p = 0.005 \)), as well as on session frequency (RERI = 2.44; 95%CI = 0.48; 4.39; \( p = 0.015 \)), was significantly higher than 0.

After stratification according to short-term regular running experience, HR were higher amongst participants with a weekly volume <2 h and those who ran <2 sessions week\(^{-1}\) in both strata. Regular running did not induce effect modification on weekly volume nor on session frequency.

The stratified analysis according to BMI revealed that the rate at which RRI occurred at any time was higher amongst the participants with a weekly volume <2 h and those who ran...

### Table 1

**Personal and sport-related characteristics of the study participants (\( n = 517 \)).**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unit/qualifier</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Years</td>
<td>42.2 ± 9.9</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td>Male 336(65.0%)</td>
</tr>
<tr>
<td>Weight</td>
<td>kg</td>
<td>71.5 ± 11.6</td>
</tr>
<tr>
<td>BMI</td>
<td>&lt;25 kg m(^{-2})</td>
<td>368(71.2%)</td>
</tr>
<tr>
<td>BMI</td>
<td>≥25 kg m(^{-2})</td>
<td>149(28.8%)</td>
</tr>
<tr>
<td>Study</td>
<td></td>
<td>Cohort 249(48.2%)</td>
</tr>
<tr>
<td>RCT</td>
<td></td>
<td>208(51.8%)</td>
</tr>
<tr>
<td>Previous injury</td>
<td></td>
<td>Yes 202(39.1%)</td>
</tr>
<tr>
<td>Regularity over the last 12 months(^{a})</td>
<td></td>
<td>No 315(60.9%)</td>
</tr>
<tr>
<td>Running experience(^{b})</td>
<td></td>
<td>Years 5(0.22%)</td>
</tr>
<tr>
<td>Regularity over the last 12 months(^{b})</td>
<td></td>
<td>No 204(39.5%)</td>
</tr>
<tr>
<td><strong>Sport-related characteristics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly running volume</td>
<td></td>
<td>&lt;2 h week(^{-1}) 259(50.1%)</td>
</tr>
<tr>
<td>Session frequency</td>
<td></td>
<td>&lt;2 sessions week(^{-1}) 258(49.9%)</td>
</tr>
<tr>
<td>Running speed</td>
<td></td>
<td>&lt;10 km h(^{-1}) 310(60%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥10 km h(^{-1}) 207(40%)</td>
</tr>
</tbody>
</table>

\(^{a}\) Three missing data.
\(^{b}\) One missing data. Descriptive data for the participants’ personal and sport-related characteristics are presented as mean (standard deviation) for continuous variables, and as counts (percentage) for categorical variables, except for running experience, for which the median and extreme values are displayed.
In period Table 526 *statistical* we *sure* that this *new* conceptual *model* was applied to the *data* and that *new* conceptual *model* was formulated *based on* the *primary* outcome. This study *aim* was *formulated* as a *new* conceptual model for RRI generation, as presented in Fig. 1. This approach does not *immediately* consider all factors as *covariates*, as suggested by established *practice*, but distinguishes between *primary* factors (training-related characteristics) and EMM (non-training-related characteristics). Furthermore, the *statistical* methods used in the *present* study are also *specific* to the *study* aim and *not* *usually* employed in the *field* of sports injury prevention. A *recent* *review* put *forward* the *great* *heterogeneity* of *statistical* methods between *studies*, which makes it difficult to *perform* the *much needed* *meta-analyses* to bring the *field* *forward*.5

The model presented here throws the basis for an *original* approach that can be adopted in future *large-scale* *prospective* *studies* and help *improve* our *understanding* of RRI *aetiology*. Rather than to *anализ* a larger *set* of training *characteristics* and potential EMM, we preferred to focus more on the *methodology* of the *analysis*. Indeed, there is virtually no *limitation* in the *number* of variables that can be tested with the *present* model.

The first step of the method applied here consists in the *crude* *analysis* of the association between *independent* primary *exposure* *variables* and RRI. This *analysis* revealed that the *groups* of runners with a *weekly volume* <2 h or a *weekly frequency* <2 displayed a higher HR. These observations are *counterintuitive*, since *common* sense would suggest the opposite, i.e. *that* a higher *weekly volume* or *frequency* would be associated with greater injury risk. *To date*, the association between *weekly running distance* and the *occurrence* of running injuries remains *unclear*. Two *high quality* *studies* reported that high *weekly mileage* (above 64 km) is a *risk factor* for lower extremity injuries.6,23 *In contrast*, higher *weekly distance* was a *strong* *protective* *factor* in *cohort* *studies*.11,29,25 It *could* be speculated that, in *habitual* recreational runners, those characterized by a higher *level* of fitness have a *decreased* risk of injury. Therefore, as suggested by others,20,30 the relationship between *weekly running volume* and RRI *risk* is *multi-dimensional* and results from a subtle combination of *overload* and *under-conditioning*. *In other words*, running experience and fitness level should be *considered* before formulating *recommendations* (e.g. *upper limits*) for *weekly volume*.

In a second phase, the size of the *effect-measure* *modification* was calculated, an *approach* rarely used in RRI *research*. *Yet*, it is *highly* recommended because it provides the *reader* with the relevant data to interpret the *effect modification* analysis.24 *Importantly*, we did identify several *assumptions* that were *significantly* influenced (positively or negatively) by *effect-measure* *modification*: previous injury or BMI. Since the effect of the *training* variables differs across the strata of these *co-variables*, it would be inappropriate to *include* them as *confounders* in the *regression* model. *Instead*, an *effect-modification* analysis is *required*, because the effect of the *confounder* is *not* *similar* across strata. This *finding* is *paramount*, and we encourage researchers in RRI research to consider analysing *effect-measure* *modification* before performing an adjusted *regression* analysis.
An example of a significant positive effect modification was the RERI = 4.69 found between weekly volume and previous injury. This means that the combined effect of running <2h week\(^{-1}\) and having had a previous injury was much worse than expected. Based on this finding, a low weekly volume and previous injury work in synergism, and it is fair to conclude that the subpopulation of individuals with low weekly volume and with previous injury are particular vulnerable to injury. Although this result may be difficult to interpret, as already discussed above, the idea here is not so much to establish a causal relationship between weekly volume and RRI, but rather to put forward the need to stratify this analysis according to previous injury. An example of a significant negative effect modification was the RERI = −2.88 found between weekly volume and BMI. Here, a lessened injury rate than expected was present for individuals with high BMI and a low running volume. In fact, a HR of 6.82 was expected based on the results from the other strata (4.70 + 2.12). Nevertheless, the HR was estimated to 2.94, and these results suggest that the subpopulation with high BMI and displaying a low weekly volume had a lessened injury rate, while the runners particularly vulnerable were those with BMI below 25 and a low weekly volume. Again, the explanations for these observations are not straightforward, and we can only speculate about the involved mechanisms. For example, it is possible that runners with a low BMI accumulate a greater mileage per running session compared to those with a high BMI, who could be more precautious and reach a given training volume through a combination of higher session frequency and lower session volume. In more general terms, the subjective perception of increased injury risk (e.g. because of a higher BMI) could lead to different behaviour and induce short-term changes in training patterns that allow for better tissue repair and a different training tolerance. To determine if these hypotheses are found, future research should be directed towards short-term changes in running routines and their relationship on cumulative tissue load, RRI and the ability for adaptive repair.\(^{16}\) Since runners generally have a fluctuating training regime, this means that methodologies taking the time-varying exposure into account are required.

Subpopulations with increased vulnerability to injury were identified in this article, which is of particular interest from a public health and injury prevention perspective. Prevention initiatives should be founded on knowledge on the causal relationship between risk factors and injury. This implies that randomized controlled trials assessing different training modalities are needed to understand the impact of training-related characteristics on RRI. In this respect, the main limitation of the present observational study is that the relationships presented here are most likely not causal. More investigations including larger numbers of runners and using controlled interventions are needed to improve our understanding of RRI aetiology. Furthermore, stratification into more subpopulations and inclusion of time-varying training-related exposures are needed to get closer to a causal pattern. Still, we believe the approach used here is “closer to causal” than the more traditional identification of risk factors using stepwise models.\(^{9–13}\)

## 5. Conclusions

The present study proposes a conceptual model in which non-training-related characteristics are considered as potential EMM, i.e. factors influencing the training load a runner is able to tolerate before injury occurs. Based on our results, we conclude that previous injury displayed a positive synergy with weekly volume and session frequency, while a negative synergy was observed between BMI and weekly volume. Future research into RRI prevention should move towards the explanation of injury mechanisms and the identification of causal relationships between training-related factors and RRI. This is a prerequisite for efficient preventive measures targeted to highest risk populations.

## 6. Practical implications

- Training-related characteristics should be considered as primary exposure of interest while non-training-related characteristics should be considered as potential EMM.
- The training load a runner is able to tolerate is affected by previous injury and BMI.
- The relationship between weekly volume or session frequency and RRI remains unclear.

## Acknowledgment

The present study was financially supported by the National Ministry of Sport, and the National Olympic Committee.

## References


