Chapter 5

The NLstart2run study: training-related factors associated with running-related injuries in novice runners

Bas Kluitenberg, Henk van der Worp, Bionka Huisstede, Fred Hartgens, Ron Diercks, Evert Verhagen, Marienke van Middelkoop

Journal of Science and Medicine in Sport; (pending minor revisions)
ABSTRACT

Objectives: The incidence of running-related injuries (RRIs) is high. Some risk factors for RRI were identified in novice runners, however, not much is known about the effect of training factors on RRI risk. Therefore, the purpose of this study was to examine the associations between training factors and RRI in novice runners, taking the time varying nature of these training-related factors into account.

Design: Prospective cohort study.

Methods: 1,696 participants completed weekly diaries on running exposure and RRIs during a 6-week running program for novice runners. Total running volume (minutes), frequency and mean intensity (Rate of Perceived Exertion) were calculated for the seven days prior to each training session. The association of these time-varying variables with RRI was determined in an extended Cox regression analysis.

Results: The results of the multifactorial analysis showed that running with a higher intensity in the previous week was associated with a higher RRI risk. Running frequency was not significantly associated with RRI, however a trend towards running three times per week being more hazardous than two times could be observed. Finally, lower running volume was associated with a higher risk of sustaining an RRI.

Conclusions: These results suggest that running more than 60 minutes spread over 2 training sessions at a lower intensity is least injurious. The finding regarding running volume is contrary to our expectations. Therefore, the findings should not be used plainly as a guideline for novices. More research is needed to unravel the person-specific training patterns that are associated with RRI.
INTRODUCTION

Running is a physical activity that is often practiced to increase health. Ironically, it also is a sports activity in which injury rates up to 94.4% are reported [1]. As a consequence, several studies have tried to identify risk factors for running-related injuries (RRIs), which is a necessary step towards development and introduction of preventive measures [2]. In spite of these attempts to clarify RRI etiology, no consistent risk factors have been identified yet. A recent systematic review on risk factors for injuries among runners concluded that a previous injury sustained in the past 12 months was the main risk factor for RRI [3]. Until now, many studies focused on intrinsic risk factors (e.g. personal, anatomical or biomechanical). However, as highlighted in a recent study, training-related factors (i.e. extrinsic risk factors) play an essential role in the injury causation model [4].

In a systematic review on the relationship between training factors and RRIs it was concluded that inconsistent findings made it impossible to identify training errors that were related to RRI [5]. The role of running volume on RRI development has been subject of many studies [6-21]. Several studies identified high running volume as a risk factor for RRIs [22, 23]. This relationship, however, was not consistently reported [20, 21, 24]. There is even evidence that increased weekly running distance is protective for knee injuries [22]. In previous studies, running volume was often averaged to one value (e.g. mean weekly running duration), subsequently its contribution to RRI was examined. This form of analysis treats running volume as a constant over time, while in reality, running volume is changing continuously. Simplifying these time-varying factors to one constant value might explain the inconsistencies reported in literature.

Nielsen et al. (2014), for example, recently conducted a study in which weekly progression in running distance was analyzed as a time-varying factor. The results showed that progression in weekly running distance with more than 30% over the previous two weeks increased the risk on specific distance-related RRIs in novice runners [25]. This study shows that time-varying training factors seem to be associated with RRIs. To unravel the relationship between training factors and RRIs, more research is needed. Therefore, the purpose of the current study was to examine the associations between training factors and RRI in novice runners, taking the time varying nature of these training-related factors into account. It was hypothesized that higher volume, intensity and frequency will increase RRI risk in novice runners.

METHODS

Study design and participants

The present study is part of the NLstart2run study, a multi-center prospective cohort study on the health effects resulting from a 6-week “Start to Run” program in novice runners [26]. All registrants of the 6-week “Start to Run” program of 2013 that is organized bi-annually (in...
March and September) by the Dutch Athletics Federation were asked to participate in the NLstart2run study (N=7,660). All registrants aged between 18 and 65 were eligible for inclusion. All participants signing the digital informed consent form and completing a baseline questionnaire were included in the study (N=1,772). Only participants completing at least one digital running diary were included in the analyses (N=1,696). The study design, procedures and informed consent procedure were approved by the Medical Ethics Committee (no. 2012/350) of the University Medical Center Groningen (UMCG), the Netherlands. The study is registered in the Netherlands Trial Registry (NTR3676).

The running program
The goal of the “Start to Run” program was to prepare participants in 6 weeks for a 20-minute run. The running program advised participants to run two to three times per week in which one training session was supervised by a licensed athletics trainer. The other training sessions had to be completed individually following a training schedule. The participants were advised to start each training session with a warm-up and finish with a cool-down both of approximately 15 minutes consisting of walking, stretching and relaxing exercises.

Data collection
A digital baseline questionnaire was sent to all participants one week before the start of the running program. This questionnaire was used to collect information on personal characteristics (gender, age, weight, height, previous running experience and previous injuries). During the running program, participants completed a weekly digital running diary. This diary was used to collect data on the number of planned training sessions and running activities (yes/no) in the preceding week. Data were collected for each training session separately. Running volume (in minutes), intensity of the training session (Ratings of perceived exertion (RPE) 6-20 point scale [27]) and the presence of pain in the lower extremities or lower back were registered for each training session. If pain was present, anatomical regions were registered with a body chart. Pain was classified as running-related pain, when the participant attributed that pain to running. For each training session with pain, a final question was used to register whether the participant was able to complete the entire training session despite of the pain. When a running diary was not completed within five days, an automatic e-mail reminder was sent. Participants were contacted by phone in case the diary was not completed one week after sending the reminder.

RRI registration
The information that was entered into the running diary was used to register RRIs during the running program. An RRI was defined as a musculoskeletal complaint in a sole body
part of the lower extremity or back which the participant attributed to running and caused a restriction in running ability (speed, distance or duration) for at least three consecutive training sessions (i.e. one week) [28]. Participants were asked not to report muscle soreness and blisters when completing the running diary.

**Analyses**

The predictors of interest were running volume in the previous 7 days, training intensity (RPE score) in the previous 7 days and number of training sessions in the previous 7 days. The previous 7 days were determined for each individual training session and, therefore, differed for each training session during the running program. The predictors of interest were calculated over the previous 7 days, excluding the current training session. The total duration (minutes) of all training sessions in the previous 7 days was determined (running volume). The number of training sessions (i.e. running frequency) and the mean intensity (mean RPE score) of all training sessions conducted in the previous 7 days were also calculated. In this manner, running volume, training intensity and running frequency were time-dependent variables, since these values differed for each training session. As a consequence of the applied method, these variables could not be calculated for running sessions within the first 7 days of the running program (i.e. were missing), because at that stage no data on the previous week were available yet. Running frequency was categorized into four categories (<2 sessions, 2 sessions, 3 sessions, >3 sessions in the previous 7 days), because a “U-shaped” pattern could be expected [5].

An extended Cox regression model was used to estimate the association of the time-dependent predictor variables with RRIs. Minutes of running were used as the time scale in the Cox models. First the independent link of the predictor variables with RRI were examined (i.e. total running volume, mean intensity and number of training sessions in the previous 7 days), following a multivariable analysis including all predictors of interest.

Training factors in the 7 days prior to an RRI might be affected by discomfort, which eventually results in an RRI. To examine whether the findings in this study are not solely the result of discomforts indicating an emerging RRI, an additional sensitivity analysis was performed in which the event of interest was not RRI, but the first report of pain. These analyses were identical to the extended Cox regression analyses as described above with the difference that the outcome was the first occurrence of pain instead of RRI.

Schoenfield residuals tests and plots were used to evaluate the proportional hazards assumption [29]. If the proportional hazards assumption was not met, the variable concerned was categorized and tested again for proportionality. To warrant valid interpretation of the results, care was taken that at least 10 RRIs per predictor variable were present in the Cox regression models [30].
Hazard ratios (HRs) with corresponding 95% confidence interval (CI) were reported for the above mentioned time-dependent predictors. HRs with $p \leq 0.05$ were considered statistically significant. All analyses were performed using R statistics (version 3.1.3) with the R survival package (version 2.38.1).

**RESULTS**

In total, 185 of the included participants sustained an RRI. However, 26 injuries occurred in the first 7 days of the running program and consequently no exposure data of the previous 7 days were available. Therefore, a total of 159 runners were included in the analyses. Demographic characteristics of all participants and for the injured and non-injured participants separately are presented in Table 1.

The assumption of proportional hazards was violated for running volume in the previous week. Therefore, running volume was categorized into four categories (<30 minutes, 30-45 minutes, 45-60 minutes and >60 minutes), which were based on the quartiles of duration. Proportionality of hazards was met for the other variables.

Table 1: Demographic characteristics of all participants and for both injured* and non-injured participants

<table>
<thead>
<tr>
<th>Variable</th>
<th>All (N=1,696)</th>
<th>Non-injured (N=1,511)</th>
<th>Injured (N=159)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>364 (21.5)</td>
<td>318 (21.0)</td>
<td>40 (25.2)</td>
</tr>
<tr>
<td>Female</td>
<td>1,332 (78.5)</td>
<td>1,193 (79.0)</td>
<td>119 (74.8)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>43.3 ± 10.0</td>
<td>43.1 ± 9.9</td>
<td>45.2 ± 10.2</td>
</tr>
<tr>
<td>BMI (kg/m2)</td>
<td>25.5 ± 4.0</td>
<td>25.4 ± 4.0</td>
<td>26.1 ± 4.5</td>
</tr>
<tr>
<td>Running experience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,023 (60.3)</td>
<td>891 (59.0)</td>
<td>113 (71.1)</td>
</tr>
<tr>
<td>Yes</td>
<td>673 (39.7)</td>
<td>620 (41.0)</td>
<td>46 (28.9)</td>
</tr>
<tr>
<td>Previous musculoskeletal complaints</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,126 (66.4)</td>
<td>1,020 (67.5)</td>
<td>95 (59.7)</td>
</tr>
<tr>
<td>Yes, not attributed to sports</td>
<td>292 (17.2)</td>
<td>245 (16.2)</td>
<td>38 (23.9)</td>
</tr>
<tr>
<td>Yes, attributed to sports</td>
<td>278 (16.4)</td>
<td>246 (16.3)</td>
<td>26 (16.4)</td>
</tr>
</tbody>
</table>

Categorical data was presented as N (%) and continuous data as means ± SD.

* Only the demographics of the injured participants who were included in the analyses were displayed. 26 injured participants were excluded from the analyses because the injury occurred during the first 7 days of running, making it impossible to calculate the predictor variables for these participants.

BMI: body mass index
Training exposures

In total 18,706 training sessions were conducted with a summed running duration of 6,732.5 hours. From these training sessions, 3,632 were conducted in the first week. Therefore no information on running volume, intensity or frequency in the previous week is available for these training sessions. Total weekly running volume was below 30 minutes in 3,830 training sessions, between 30 to 45 minutes in 4,068 training sessions, between 45 to 60 minutes in 2,998 training sessions and above 60 minutes in 4,178 training sessions. Running frequency in the previous 7 days was below 2 in 2,214 training sessions, 2 in 6,453 training sessions, 3 in 6,036 training sessions and above 3 in 371 training sessions. Mean training intensity was 12.1 ± 1.5 on the RPE scale.

Training-related risk factors

The results of the univariable and multivariable Cox regression analyses are shown in Table 2. Higher intensity was associated with RRI occurrence in both the univariable (HR 1.29; 95% CI 1.18-1.41) and multivariable analyses (HR 1.28; 95% CI 1.18-1.40). Running more than 60 minutes in the previous 7 days was protective for the occurrence of an RRI in both the univariable (HR 0.52; 95% CI 0.29-0.93) and multivariable analysis (HR 0.41; 95% CI 0.20-0.86). Running frequency was not associated with RRI in both the univariable and multivariable analyses.

Table 2: Results from the univariable and multivariable Cox regression analyses in which running volume, intensity (RPE score) and frequency in the previous 7 days were predictors for RRI.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariable</th>
<th>Multivariable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>Running volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 30 minutes</td>
<td>1 (reference)</td>
<td>0.164</td>
</tr>
<tr>
<td>30-45 minutes</td>
<td>0.73 (0.48-1.13)</td>
<td>0.188</td>
</tr>
<tr>
<td>45-60 minutes</td>
<td>0.71 (0.42-1.18)</td>
<td>0.028</td>
</tr>
<tr>
<td>&gt; 60 minutes</td>
<td>0.52 (0.29-0.93)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Intensity</td>
<td>1.29 (1.18-1.41)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2 sessions</td>
<td>1.16 (0.74-1.81)</td>
<td>0.520</td>
</tr>
<tr>
<td>2 sessions</td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td>3 sessions</td>
<td>1.20 (0.84-1.72)</td>
<td>0.307</td>
</tr>
<tr>
<td>&gt; 3 sessions</td>
<td>1.71 (0.51-5.71)</td>
<td>0.386</td>
</tr>
</tbody>
</table>

Sensitivity analysis on first pain occurrence

The results of the sensitivity analysis on first pain occurrence are presented in Table 3. The results from the sensitivity analyses showed a similar pattern as the normal analyses on RRI.
First pain occurrence was associated with running less than 30 minutes in the previous week (HR 0.67; 95% CI 0.55-0.83, HR 0.53; 95% CI 0.41-0.68, HR 0.45; 95% CI 0.34-0.59).

Table 3: Results from the sensitivity analyses, examining running volume, intensity (RPE score) and frequency in the previous 7 days as predictor for occurrence of first pain.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariable HR (95% CI)</th>
<th>Univariable p-value</th>
<th>Multivariable HR (95% CI)</th>
<th>Multivariable p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Running volume</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 30 minutes</td>
<td>1 (reference)</td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td>30-45 minutes</td>
<td>0.67 (0.55-0.83)</td>
<td>&lt;0.001</td>
<td>0.63 (0.50-0.79)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>45-60 minutes</td>
<td>0.53 (0.41-0.68)</td>
<td>&lt;0.001</td>
<td>0.45 (0.33-0.61)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&gt; 60 minutes</td>
<td>0.45 (0.34-0.59)</td>
<td>&lt;0.001</td>
<td>0.38 (0.27-0.53)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Intensity</td>
<td>1.03 (0.98-1.09)</td>
<td>0.271</td>
<td>1.02 (0.97-1.08)</td>
<td>0.432</td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2 sessions</td>
<td>1.24 (1.02-1.51)</td>
<td>0.033</td>
<td>0.85 (0.67-1.08)</td>
<td>0.183</td>
</tr>
<tr>
<td>2 sessions</td>
<td>1 (reference)</td>
<td></td>
<td>1 (reference)</td>
<td></td>
</tr>
<tr>
<td>3 sessions</td>
<td>1.01 (0.85-1.20)</td>
<td>0.904</td>
<td>1.18 (0.98-1.42)</td>
<td>0.084</td>
</tr>
<tr>
<td>&gt; 3 sessions</td>
<td>0.73 (0.39-1.34)</td>
<td>0.308</td>
<td>0.89 (0.47-1.67)</td>
<td>0.719</td>
</tr>
</tbody>
</table>

**DISCUSSION**

This is the first study in which the association of running volume, frequency and intensity with RRI was examined while taking into account the time-varying nature of these variables in novice runners. The results from the present study showed that running at a higher intensity increased RRI risk, whereas running frequency was not associated with RRI in novice runners. The results of this study also suggest that in this population of runners, running more than 60 minutes per week is protective for sustaining an RRI compared to running less than 30 minutes per week. This is an unexpected finding and it should be stressed that this should not be interpreted as a green light for novice runners to increase running volume above 60 minutes per week.

Surprisingly, the association between perceived running intensity and RRI was not studied yet. The results from our study suggest that training at a higher intensity (higher RPE score) is associated with a higher risk on RRI. Some studies did examine the link between mean running speed and RRI risk [11, 12, 15, 16, 21, 24]. These studies reported conflicting results, presumably due to the fact that running speed is not equal to perceived running intensity and mean running speed was kept constant over time [5]. Based on the results of the present study, novice runners should be advised not to run each training session on a too high intensity.

Running frequency was not significantly associated with RRI. However, a trend was observed towards running three times per week being more hazardous than two times. The
multivariable analysis suggests that participants who run more than 60 minutes in the previous week spread over two training sessions had the lowest injury risk. Several studies examined the relationship between running frequency and injury risk. In contrast to our results, most studies suggested a “U-shaped” pattern between running frequency and RRI risk [5]. It should be noted, however, that in the current study a population of novice runners was studied while most studies did follow other populations of runners. This might explain these differences.

Higher running volume is often identified as risk factor for RRI [6-8, 11-15, 17-19, 31]. Interestingly, the results from the current study show an opposite trend. The two main differences between this study and previous studies are the method of analysis and the population of runners followed. In this study, weekly running volume was determined for each week separately, which fluctuated for each training session, while most previous studies asked runners to report their normal weekly running volume. In reality, however, a runners’ running volume depends on many factors and is not constant over time. Statistical models that allow this exposure to vary over time, therefore, reflect reality better. A limitation of using such models in observational studies is that the reason for differences over time is not always clear. For instance, it is plausible that an RRI is preceded by physical complaints and as a result the participant decides to shorten or skip a training session (i.e. reduced running volume). Consequently, a reduced running volume might be associated with an increased injury risk. To ensure that this phenomenon did not influence our findings, a sensitivity analysis was conducted on the occurrence of the first physical complaint that was reported by the runners. This sensitivity analysis showed that a higher weekly running volume was also associated with a decreased risk on physical complaints. This indicates that early physical complaints preceding an RRI did not explain our findings.

A recent study among recreational runners showed that participants with a running volume below 2 hours per week had a higher risk compared to runners with a weekly running volume above 2 hours [4]. Even though this study did not take the time-varying nature of running volume into account, the results were comparable to ours. It is often assumed that little running experience is associated with a higher risk on RRI [22]. It is likely that runners with more running experience have a higher weekly running volume. This might explain why lower running volume was related to RRI in recreational runners [4]. In the current study, however, this effect is expected to be small, because all runners were participating in a “Start to Run” program and therefore had little to no running experience. It is also possible that participants with specific characteristics pick up running more easily. Possibly, these runners had a higher running volume and at the same time had a lower risk of RRI. Future well controlled cohort studies or randomized controlled trials are necessary to clarify the association between running volume and RRI taking into account personal characteristics.
Limitations

In the current study we did not account for personal characteristics, because it is incorrect to include such factors as confounders. Personal characteristics certainly affect RRI risk, however, it is important to realize that these factors influence RRI risk through training [4]. Personal characteristics do not directly impact RRI (without training no RRI will occur), but these factors modify the response to training. Personal characteristics should therefore be handled as effect measure modifiers [4]. As a result, analysis should be stratified on personal characteristics when accounting for these factors. Unfortunately it was not justified to do this kind of analysis in the current study, because the number of RRIs in each stratum turned out to be too low (below 5 RRIs).

We made the assumption that training characteristics in the seven days prior to RRI were predictive for occurrence of the RRI. It might be possible that this time span was too small and we should have looked back further. We believe, however, that this method was most suitable, because of the short duration and the high number of injuries in the first few weeks of the running program [33].

In this study, we did not take differences between weeks into account. It can be argued that changes in training regime might be related to RRI though [25]. Due to the short duration of the running program it was not possible to analyze these differences. Future studies are therefore advised to take changes in running exposures into account, preferably in well controlled cohort studies or randomized controlled trials.

All runners were participating in a running program with a predefined training schedule. Therefore, it could be argued that not enough variation in training characteristics was present. During the running program, one training session was supervised and the other sessions were done individually. This resulted in many differences, allowing the analyses performed.

CONCLUSION

The results of this study showed that running at a higher intensity increased RRI risk in novice runners and running frequency was not significantly associated with RRI. A higher running volume turned out to be protective for RRI. This result is contrary to our expectations and we do not have a uniform explanation for this finding. It should be stressed that novice runners are not necessarily advised to start running with a volume above 60 minutes though. Future well controlled studies should examine the effect of training characteristics to RRI taking into account the time-varying nature of these variables and treat personal characteristics as effect measure modifiers.
ACKNOWLEDGEMENTS

The NLstart2run study was funded by a grant of ZON-MW (50–50305-98-12001). The authors wish to thank all runners who participated in this study, the Dutch Athletics Federation for providing the opportunity to include participants from the Start to Run program, and Astrid de Vries, Corien Plaggenmarsch and Saskia van de Zande for their assistance during the data collection.

REFERENCES

The NLstart2run study: training-related factors associated with running-related injuries in novice runners