Risk models for lower extremity injuries among short- and long distance runners: A prospective cohort study

Running head: risk model for running injuries

Authors: Dennis van Poppel, PT MSc[1], Gwendolijne GM Scholten-Peeters, PT, PhD[1,2], Marienke van Middelkoop PhD[3], Bart W Koes[3] Prof, Arianne P Verhagen PhD[1,3]

[1] Avans University of Applied Sciences, Research Group Diagnostics, Breda, The Netherlands

[2] Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, The Netherlands,

[3] Erasmus medical center University. Dept General Practice, Rotterdam, The Netherlands

Funding: no funding

Conflict of interest: none of the authors had conflicts of interest

All authors had access to the data and a role in writing the manuscript

Article type: clinical research study

Word count: abstract: 248; text: 2939; nr of references: 30

Key Words: running-related injuries, risk factors, running, risk model

Correspondence to:

Dennis van Poppel, MSc

Avans University of Applied Sciences, Research Group Diagnostics, PO Box 90116, 4800 RA Breda, The Netherlands. **Email: d.vanpoppel@avans.nl**

ABSTRACT

Background: Running injuries are very common. Risk factors for running injuries are not consistently described across studies and do not differentiate between runners of long- and short distances within one cohort.

Objectives: The aim of this study is to determine risk factors for running injuries in recreational long- and short distance runners separately.

Design: A prospective cohort study.

Methods: Recreational runners from four different running events are invited to participate. They filled in a baseline questionnaire assessing possible risk factors about 4 weeks before the run and one a week after the run assessing running injuries. Using logistic regression we developed an overall risk model and separate risk models based on the running distance.

Results: In total 3768 runners participated in this study. The overall risk model contained 4 risk factors: previous injuries (OR 3.7) and running distance during the event (OR 1.3) increased the risk of a running injury whereas older age (OR 0.99) and more training kilometers per week (OR 0.99) showed a decrease. Models between short- and long distance runners did not differ significantly. Previous injuries increased the risk of a running injury in all models, while more training kilometers per week decreased this risk.

Commented [AV1]: Were protective is duidelijker denk ik

Conclusions: We found that risk factors for running injuries were not related to running distances. Previous injury is the most important generic risk factor for running injuries, as is a weekly training distance. Prevention of running injuries is important and a higher weekly training volume seems to prevent injuries to a certain extent.

Keywords: running, injuries, running related injuries, risk models

INTRODUCTION

Running is an increasingly popular form of physical activity in Western countries.[1,2] In 2008, about 11.5% of the population in the US ran, and 3.4% of this group ran two times a week or more on average.[3] Between 2000 and 2010 the number of half marathon runners in Switzerland increased from 2904 to 8690 female runners and from 9333 to 21583 male runners.[2] In 2012 almost 2 million Dutch people participated in running activities.[4] This is about 11% of the total Dutch population. Although several health benefits are attributed to running activities,[5,6] injuries also occur frequently.[7-10] In the Netherlands about 32% of the runners get injured each year.[4] Most running injuries occur in the lower extremities[11-14] with an incidence varying from 19.4 to 79.3%.[1] This wide variation in incidence is likely due to differences in study-populations and definition of injuries.[8] The most common site of running injuries is the knee.[13-15]

Several studies evaluated risk factors for running injuries.[12,16] The most important risk factors found are: a history of previous injuries and an increased training volume per week in male runners.[1,10,16] The common belief is that factors like body mass index (BMI), running experiences, types of shoes and training characteristics (duration, frequency of running, training distance, running speed, warm-up and exercise habits before running) are also associated with increased risk of running injuries but no statistically significant association has been found yet.[7,9,13,17,18] This may be due to the fact that most research on risk factors for running injuries has been performed in homo- and heterogeneous groups of runners, varying from military personnel to recreational runners, running 5 km to marathon distances (42,195km).[1,18] Training related characteristics such as volume, frequency, duration and intensity of training differ between runners of different distances.[19] Half marathon runners had, compared to marathon runners, significantly less running experience (7.9 years versus 10.5 years), run less weekly training kilometers (minimum weekly distance 16.2 to 45.2 km versus 22.8 to 63.3 km), and run less weekly running hours (3.9 versus 4.8 hrs).[19] Some gender-specific risk factors were also found.[16] Overall, women are at lower risk of developing running related injuries.[16] Previous injuries, running experience (0-2 years), restarting running and having a weekly running distance of more than 40 miles are associated with greater injury risk in men than in women. Age, previous sports, running on concrete surface, participating in marathons, weekly running distance (30-39 miles), and wearing running shoes for 4-6 months were associated with an increased risk of running injuries in females than in males.[16] More females started running, mainly 10 km and half marathons, and the male/female ratio changed from 3:2 to 2:5.[2] In general, risk factors vary between different studies as the result of heterogeneity of the study population, definition of injury, type of runners (recreational or elite) and running distance.[8,14]

No previous studies prospectively evaluated the incidence of running injuries and possible different risk factors for running injuries in recreational short- and long distance runners. Therefore, the aim of this study is to assess the risk factors for running injuries among recreational runners on several running distances during the race and determine whether risk factors differ between the various distances.

METHODS

Design. A prospective cohort study with a 12-month follow-up. Runners were invited to participate in the study and were followed-up for 12 months by using web-based questionnaires. The Medical Ethical Committee of the Erasmus Medical Centre (MEC-2009-319) approved this study.

Study participants. Participants (>18 years) of four different yearly national running events in The Netherlands were invited. These running events were the Amgen Singelloop Breda (twice: October 2009 and October 2011), ABN AMRO Marathon Rotterdam (April 2012), and the Lage Landen Marathon Eindhoven (October 2012). The runners could run a variety of distances including the marathon (42,195 km), half marathon (21,095 km), 15 km, 10 km and 5 km runs. Since there was a low turnout on the 15 km distance, these runners were combined with the 10 km group, forming a moderate distance group: short distance (5 km), moderate distance (10 and 15 km), half marathon and marathon.

Participants were invited if they subscribed digitally as individual recreational runners at least 4 weeks before the start of the running event and provided a valid email address. Excluded were competition and business runners.

Procedure. Participants received information via email about the study accompanied by a link to an online baseline questionnaire,~~. W~~which was developed and used previously.[15,17,20] All participants who returned the baseline questionnaire and agreed with the informed consent, were included in the study and received a follow-up questionnaire one week after the event (and 3, 6, 9 and 12 months after the event). Non-responders received a reminder within one week. For this manuscript we only use the baseline data and the data of one week after the event.

Baseline determinants. At baseline, runners were asked to complete questions about a) sociodemographic characteristics (e.g. age, gender, height, weight, education, lifestyle (e.g. smoking, alcohol)), b) training related characteristics (e.g. type of training, weekly training frequency, weekly running distance) and c) other running related risk factors, based on the literature (e.g. years of running experience, running terrain, and previous running injuries during the last year).

Categorical determinants with the answer options: always, often, sometimes, rarely, or never, were dichotomized into 'often' (always, often) and 'sometimes' (sometimes, rarely, never), in accordance with a previous study.[17] BMI was calculated based on height and weight and included in the analysis as a continuous variable. The variable 'previous injuries in 12-months preceding the event' was dichotomous (yes/no).

A priori we defined 22 determinants relevant for the analysis: age, gender (male/female), BMI, alcohol use (yes/no), daily smoking (yes/no), education level (high/low), specific feeding supplements (yes/no), injuries in the previous 12 months (yes/no), participation in an organized running group (yes/no), running experience (years), training on firm underground (yes/no), weekly training hours, frequency and kilometers, average running speed, long distance training, interval training (yes/no), stretching before and after the training (yes/no), warming up before and after the training (yes/no) and running distance in the event (5km, 10/15km, half marathon or marathon).

Follow-up measurement. The follow-up questionnaire (one week after the event) obtained information regarding the running event itself (running distance and performance), new running injuries during these events, location of injuries, and pain intensity measured with an 11-point Numeric Rating Scale (NRS).[21,22]

Outcome The outcome of interest was the presence of new running injuries during the running events as reported the one-week follow-up. Running injuries were defined as self-reported complaints of muscles, joints, tendons or bones in the lower extremity (hip, groin, thigh, knee, lower leg, ankle, foot and toe) due to running activities by which the running intensity or frequency was reduced, or medical consultation was needed.[7,13,17,23]

Statistical analysis.

Descriptive analysis. If participants subscribed to more than one of the running events (e.g. Singelloop 2009 and 2012), we only included the data of the first running event in which the participant took part. We calculated descriptive statistics (frequencies) for baseline characteristics, including means and standard deviations. In case the data did not show a normal distribution, we presented medians and interquartile ranges. We used the Independent Samples T-test to analyze differences between responders and non-responders.

Risk model development. Before developing a multivariate logistic regression model we evaluated multicollinearity between potential determinants; if a correlation between two determinants was ≥ 0.8 only one of the determinants was chosen for the multivariate analyses. First, the multivariate analysis was performed in the total cohort (method Backward Wald, p<0.1 for exclusion). Secondly, we calculated risk models for each distance separately. Results were expressed in Odds Ratios (ORs). In case of missing variables, participants were excluded from the multivariate analysis. We complied with the 1 in 10 rule (one determinant per every 10 injuries) in the analysis, and selected the appropriate number of determinants a priori, based on the literature.[25]

Potential risk predictors. An overview of all 22 determinants is given in Table 1. For the 5km runners we could enter 5 to 6 variables in the regression model. We choose to enter the variables that were found relevant in a previous study (age, previous injury, weekly training distance, interval training and participation in organized running groups).[13] Among 10-15km runners 21 (all except running distance) variables could be entered into the regression analysis. Finally, we included 18 determinants in the analysis of the half marathon group (age, gender, BMI, alcohol use, daily smoking, education level, specific feeding supplements, injuries in the previous 12 months, participation in an organized running group, running experience, training on firm underground, weekly training hours, frequency and kilometers, average running speed, long distance training, interval training). The same determinants were used in the analysis for the marathon runners.

Model performance. Lastly, performance measures of the model were calculated: explained variance (R^2) and the area under the curve (AUC)). The AUC represents the ability of the risk model to distinguish between patients with or without an injury at the 1 week follow-up and ranges from 0.5 (no discrimination) to 0.1 (perfect discrimination).[25] An AUC ≥ 0.7 is considered good discrimination and an AUC between 0.6 and 0.7 as moderate discrimination.

Commented [AV2]: Waarom hier een passieve formulering?

Data were analyzed using the Statistical Package for Social Sciences (SPSS version 23, Inc, Chicago, Illinois).

Construction of the nomogram. To make the model suitable for use in clinical practice, we transformed the regression equation into a nomogram or score chart. The coefficients in the regression equation were multiplied by 15 and rounded to the nearest integer to obtain the score per predictor. Multiplication by 15 was chosen to get the majority of the coefficients close to an integer, thereby minimizing the effects of rounding. The sum of all scores reflects the probability of getting an injury during a running event.

RESULTS

Participants. In total 17,891 participants received an invitation to participate by email, of which 3,768 runners (21.1%) returned the baseline questionnaire. In total 383 participants ran 5km, 1,189 participants ran 10km, 185 ran 15km, 927 participants ran the half marathon and 1,055 participants the marathon. Added numbers do not match up completely because of some missings.

Baseline. The mean age of the runners was 42.8 years, with a range from 16–83 years; 60.8% were male and the average BMI was 23.4 (see Table 1). The percentage of males was highest in the marathon group (78.5%) and lowest in the 5km group (23.2%). Also the percentage of runners using food supplements was highest in the marathon group (52.9%) and lowest in the 5km group (8.6%). Almost half of the runners replied with a "yes" when asked whether they had suffered running injuries during the 12 months before the baseline questionnaire.

Insert Table 1, please

Follow-up. At the follow up (one week after the event) in total 2,763 runners (73.3%) responded to the follow-up questionnaire (see Figure 1). We found statistically significant differences between responders and non-responders at follow-up for some variables. Non-responders were notably younger, had a higher BMI, ran shorter distances more often and there were more female responders compared to the rest of the group (see table 2). Although statistically significant, the differences between the groups were small.

Insert Figure 1 and Table 2, please

In total 2,566 participants (92.9%) started and finished, 46 participants did not finish, and 151 persons did not start due to sickness or injuries. Of 2,721 runners we received data on injuries incurred between answering the baseline questionnaire and the follow-up (i.e. either since the baseline questionnaire but before the event or during the event). Overall, 811 runners (21.5 %) reported one or more running injuries at the follow-up; 5km: 17.5% (67/250), 10-15km: 18.7% (257/981), half marathon: 23.1% (214/708) and marathon: 25.2% (266/762).

Risk models

Total cohort. In total 2,369 runners were included in the multivariable analysis, of which 709 (out of 811) had a running injury. We found no correlations between determinants above 69%, so no determinants were removed from multivariable regression analysis.
Multivariable regression analysis resulted in a risk model including 4 determinants (see table 3): two of which were risk factors (increasing the risk of an injury): previous injuries (OR 3.7; β 1.30) and running distance during the event (OR 1.3; β 0.27), two others were protective: older age (OR 0.99; β -0.013) and more training kilometers per week (OR 0.99; β 0.012). The Hosmer & Lemeshow test is not significant, indicating a good fit of the model. The overall risk model has an explained variance (Nagelkerke's R^2) of 12%, AUC of 68.4% (66.2–70.6), and it correctly classifies 70% of the runners.

Insert Table 3, please

Analyses per running distance. Since the running distance was a statistical significant risk factor we also calculated a risk model per running distance (see table 3). We found a 5km risk model including 4 determinants: age (OR 0.97: β -0.026), previous injury (OR 4.1: β 1.400) and weekly training distance (0.95, β -0.057). Among 10-15km runners we found a 10km risk model including 5 determinants: age (OR 0.98; β -0.018), BMI (1.1; β 0.074), previous injury (OR 3.8; β 1.325), weekly training distance (0.97; β -0.026) and training frequency (OR 1.3; β 0.279) which correctly classified 72.7% of the runners (R^2 = 13.4%). For the half marathon and marathon runners, the regression analysis revealed a model including 2 determinants: previous injuries (OR 3.3; β 1.204 half marathon runners and OR 4.3; β 1.448 in marathon runners) and weekly training distance (OR 0.98; β -0.013 in both risk models). For all risk models the Hosmer & Lemeshow test was not significant, indicating a good fit and all risk models correctly classify 66-76% of the runners. Furthermore, the AUC for all risk models was moderate.

Commented [AV3]: Deze mag je ook presenteren, maar die heb ik niet gebruikt voor het nomogram

Nomogram

The nomogram that we derived from the logistic regression model is presented in Table 4. The weight of an item is based on its β coefficient in the logistic regression equation. Table 4 also provides the score chart legend to convert the total score into the predicted probability of persistent complaints.

Discussion

We found an incidence of running injuries between 17.5% (5km) and 25.2% (marathon) depending on the running distance. Running distance during the event appeared to be a statistical significant risk factor for developing running injuries. The distance specific risk models were quite comparable; two factors were present in all risk models: previous injury increased the risk of running injuries and higher number of weekly training kilometers decreased the risk.

Comparison with other studies. For the marathon the incidence of running injuries is in line with previous studies among marathon runners.[15,26] This is the first study that developed risk models for running injuries across different running distances in one cohort. Our hypothesis that risk factors for running related injuries vary, depending on the running distances, seems to not be confirmed. We rather found comparable distance specific risk models.

A review described that lower age is a protective factor and older age is associated with an increased risk for running injuries.[16] A possible explanation for our contradictory finding could be that relatively older runners are fitter or better prepared than younger ones. Probably, if they would have had running injuries earlier, they would have stopped their running activities (healthy volunteer bias).[26] Also, knowledge of their body could be better than in younger runners so overuse is less likely to appear.[26] Another explanation could be that peak ground reaction forces (GRF) in older runners seem to be lower than in younger runners and therefore they may be at lower risk. When GRF are higher, loading of joints and muscles is increased and possibly overuse injuries are less likely to appear.[27]

In this study age was only included in the final risk models for the shorter distances and in the overall risk model. Older age was a significant protective factor although odds ratios are small (OR 0.97-0.99) This is due to the fact that age is a continuous variable. Nevertheless it contributes statistically to the whole risk model.

Gender was not included in any of the risk models; which is in contrast with a recent systematic review,[16] which showed that male gender is a risk factor for running injuries. However, a recent cohort study showed that female recreational runners have a different type of knee loading in comparison to males; which could explain differences in injury rates.[28]

Strengths and limitations. Strength of this study is the large population of runners included. Moreover, no previous studies have assessed risk factors in one cohort in four different distances. This study also has some limitations. One of the limitations is the diagnosis of running injuries since we used the self-reported complaints definition.[17,23] There was no physical examination in this study to objectify an injury. Also, participants might have applied the criteria for an injury differently in answering the questions. This could have led to an overestimation of running related injuries because complaints of muscle soreness could be interpreted as an injury according to our definition. On the other hand, there could also be an underestimation while participants did not report any injuries because of the absence of impairments in training or competition and/or medical consultation in regard to the definition from the recent consensus.[29]

Another limitation is that all determinants were obtained by self-reported questionnaires and the validity of the questionnaire is unknown. Therefore, it is possible that we have missed potential relevant risk factors such as psychosocial factors. Self-report studies are inherently biased by the person's feelings at the time they filled out the questionnaire.[30]

Despite these limitations, the results of this study may contribute to the growing body of knowledge of running injuries, especially at other distances than marathon runners only.

CONCLUSION

We found that risk models for short- and long distance runners did not differ much. Previous injury is the most important generic risk factor for running injuries, as is weekly training distance. To prevent running injuries three risk factors seem to be important: age, previous injuries and weekly training volume. Previous injuries cannot be modified, although it became clear that it is important to prevent running injuries as this factor majorly contributes to the risk models. Runners should pay attention to their weekly training volume, as a higher weekly volume seems to be protective. There might be an optimum weekly training volume (per running distance of the event), but we were unable to assess that. Future research might also consider individual athletes' relative changes in training loads or the training load compared to the distance ran, rather than the absolute load.

REFERENCES

1. Van Gent RN, Siem D, Van Middelkoop M, Van Os AG, Bierma-Zeinstra BW, Koes BW. Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. Br J Sports Med 2007;41:469-480.
2. Aschmann A, Knechtle B, Cribari M, Rüst CA, Onywera V, Rosemann T, Lepers R. Performance and age of African and non-African runners in half- and full marathons held in Switzerland, 2000-2010. Open Access J Sports Med. 2013;4:183-92.
3. Messier SP, Legault C, Schoenlank CR, Newman JJ, Martin DF, Devita P. Risk factors and mechanisms of knee injury in runners. Med Sci Sports Exerc 2008;40:1873-1879.
4. Van Hespen A, Stubbe J, Stege S, Ooijendijk W. Injury free running? Leiden: TNO KvL; Injury information system 2012;1986-2012, VeiligheidNL
5. Taunton JE, Ryan MB, Clement DB, McKenzie DC, Lloyd-Smith DR, Zumbo BD. A prospective study of running injuries: the Vancouver Sun Run "In Training" clinics. Br J Sports Med 2003;37:239-244.
6. Verhagen E. Prevention of running-related injuries in novice runners: are we running on empty? Br J Sports Med. 2012;46(12):836-7.
7. Van Poppel D, Koning, J. de, Verhagen. A.P., Scholten-Peeters, G.G.M., Risk factors for lower extremity injuries among half marathon and marathon runners of the Lage Landen Marathon Eindhoven 2012. A prospective cohort study in the Netherlands, Scand J Med Sci Sports; 2016;26(2):226-34
8. Kluitenberg B, Middelkoop M, Diercks R, van der Worp H. What are the difference is injury proportions between different population of runners? A systematic review and meta-analysis. Sports Med 2015: 2015;45(8):1143-61
9. Malisoux L, Nielsen RO, Urhausen A, Theisen D. A step towards understanding the mechanisms of running-related injuries. J Sci Med Sport. 2014;S1440-2440(14)00140-6.
10. Saragiotto BT, Yamato TP, Hespanhol Junior LC, et al. What are the main risk factors for running-related injuries? Sports Med. 2014;44:1153–63.
11. Chang WL, Shih YF, Chen WY. Running injuries and associated factors in participants of ING Taipei Marathon. Phys Ther Sport 2012;13:170-174.

12. Lopes AD, Hespanhol Júnior LC, Yeung SS, Costa LO. What are the main running-related musculoskeletal injuries? A Systematic Review. Sports Med 2012;42:891-905.
13. Van Poppel D, Scholten-Peeters GGM, Van Middelkoop M, & Verhagen AP. Prevalence, incidence and course of lower extremity injuries in runners during a 12-month follow-up period. Scan J Med Sci Sports; 2014;24(6):943-9.
14. Kluitenberg B, van Middelkoop, Smits DW, Verhagen E, Hartgens, F, Diercks R, Worp, H. van der, The NLstart2run study: Incidence and risk factors of running-related injuries in novice runners, Scand J Med Sci Sports 2015: volume: e515-23
15. Van Middelkoop M, Kolkman J, Van Ochten J, Bierma-Zeinstra SM, Koes BW. Prevalence and incidence of lower extremity injuries in male marathon runners. Scan J Med Sci Sports 2008;18:140-144.
16. van der Worp Maarten P, Ten Haaf Dominique SM, van Cingel Robert, de Weijer Anton, Nijhuis van der Sanden Maria WG, Staal Bart. Injuries in Runners; A systematic review on Risk Factors and Sex Differences. PLoS One. 2015;10(2): e0114937.
17. Van Middelkoop M, Kolkman J, Van Ochten J, Bierma-Zeinstra SM, Koes BW. Risk factors for lower extremity injuries among male marathon runners. Scan J Med Sci Sports 2008;18:691-697.
18. Knapik JJ, Orr R, Pope R, Grier T. Injuries And Footwear (Part 2): Minimalist Running Shoes. J Spec Oper Med. 2016 Spring;16(1):89-96.
19. Ristolainen L, Heinonen A, Turunen H, Mannström H, Waller B, Kettunen JA, Kujala UM.Type of sport is related to injury profile: A study on cross country skiers, swimmers, long-distance runners and soccer players. A retrospective 12 month study. Scan J Med Sci Sports 2010;20:384-393.
20. Zillmann T, Knechtle B, Rust CA, Knechtle P, Rosemann T, Lepers R. Comparison of training and anthropometric characteristics between recreational male half-marathoners and marathoners. Chin J Physiol 2013;56(3):138-46
21. Van Middelkoop M, Kolkman J, Van Ochten J, Bierma-Zeinstra SM, Koes BW. Course and predicting factors of lower-extremity injuries after running a marathon. Clin J Sport Med 200;17(1):25-30
22. Gallasch CH, Alexandre NM. The measurement of musculoskeletal pain intensity: a comparison of four methods. Rev Gaúcha Enferm 2007;28:260-265.

23. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. J Shoulder Elbow Surg 2009;18:920-926.
24. Macera CA, Pate RR, Powell KE, Jackson KL, Kendrick JS, Craven TE. Predicting lower-extremity injuries among habitual runners. Arch Intern Med 1989;149:2565-2568.
25. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373-1379.
26. Rubin, D.E. (1987). Multiple imputation for nonresponse in surveys. New York: Wiley
27. Sattertwhaite P, Norton R, Larmer P, Robinson E. Risk factors for injuries and other health problems sustained in a marathon. Br J Sports Med 1999;33:22-26.
28. Kline Paul W, PT[1] and D.S. Blaise Williams, III. Effects of normal aging on lower extremity loading and coordination during running in males and females. Int J Sports Phys Ther. 2015;10(6):901–909.
29. Sinclair J, Selfe J: Sex differences in knee loading in recreational runners. J Biomech. 2015;48(10):2171-5.
30. Timpka T, Alonso JM, Jacobsson J, et al. Consensus statement on injury and illness definitions and data collection procedures on epidemiological studies in Athletics (track and field). Br J Sports Med 2014;48:483–490.
31. Schwarz N. Self-reports: How the question shape the answers. Am Psychol 1999;54:93-105.

Legend

Figure 1: Flow chart participant

Table 1: Characteristics of the running cohorts

Table 2: Characteristic of responders versus non-responders

Table 3. Multivariate regression models (backward wald) for running injuries

Figure 1: Flow chart participant

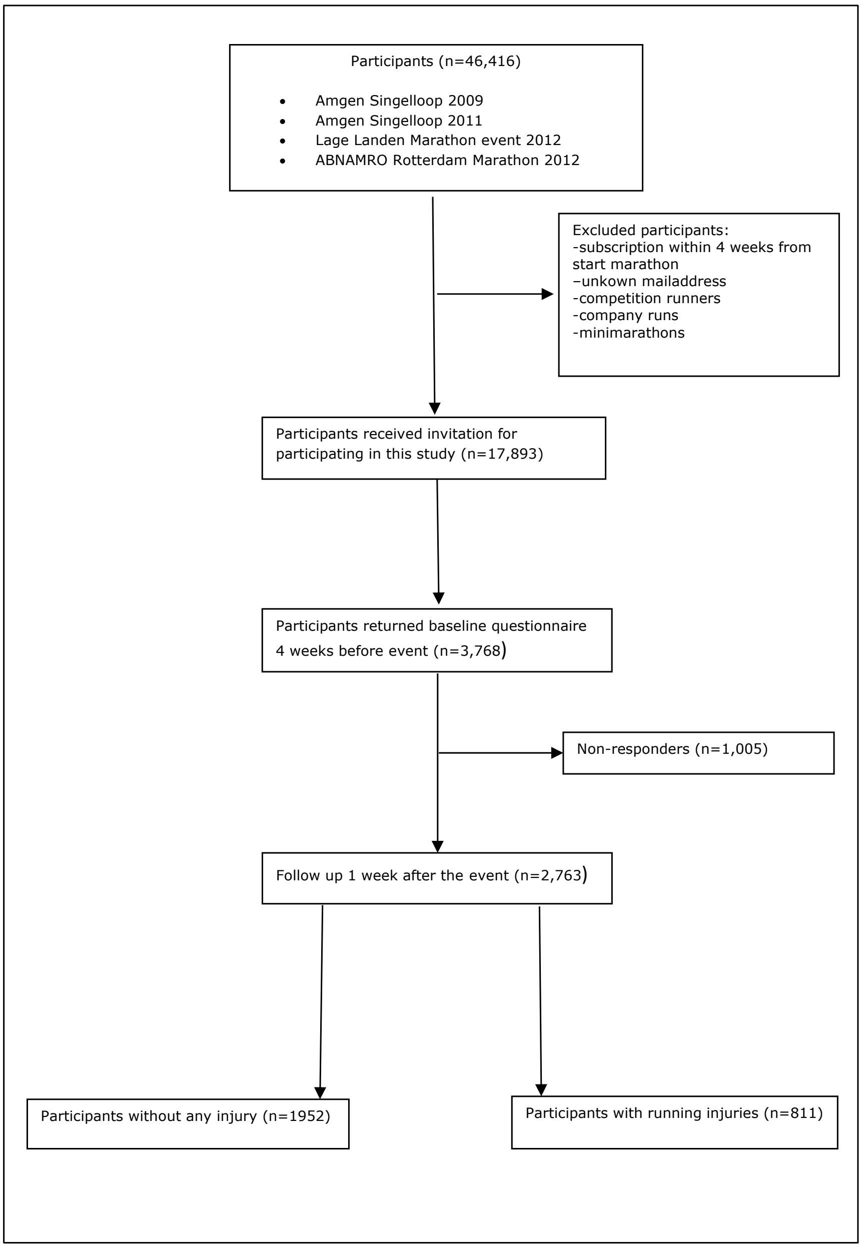

Table 1: Characteristics of the running cohorts

Determinants*	5 km n = 383	10-15 km n = 1374	Half marathon n = 927	Marathon n = 1055	Total n =3768
Demographic determinants					
Gender: males (%)#	89 (23.5)	695 (50.6)	642 (69.3)	828 (78.5)	2270 (60.2)
Age in years, mean (SD), range #	39.1 (12.4), 16-73	41.8 (11.4), 16-77	43.2 (11.4), 17-75	45 (9.6), 19-83	42.8 (11.2), 16-83
BMI, mean (SD)#	23.8 (3.1)	23.6 (2.6)	23.2 (2.4)	23.1 (2.2)	23.4 (2.5)
Education level, higher education (%)	300 (78.3)	1045 (76.1)	716 (77.2)	795 (75.4)	2857 (76.3)
Daily smoking: yes (%)	291 (76)	60 (4.4)	38 (4.1)	32 (3.0)	161 (4.3)
Alcohol use: yes (%)	29 (7.6)	1152 (83.8)	725 (82.5)	847 (80.3)	3080 (81.7)
Special feeding supplements: yes (%)#	33 (8.6)	163 (11.9)	218 (23.5)	558 (52.9)	979 (26.0)
Previous injury 12 months: yes (%)#	175 (45.7)	536 (45.0%)	520 (56.1)	626 (59.3)	1976 (52.2)
Training related determinants					
Trainings distance, km/week, mean (SD), range #	12 (7), 2-50	20 (11.2), 1-81	31.7 (14.4), 1-87	46.5 (17.6), 1-100	29.5 (18.4) 1-100
Training frequency, times/week, mean (SD) range #	2.3 (0.7), 1-6	2.4 (0.8), 1-12	2.9 (0.9), 1-7	3.7 (1.1), 1-12	2.9 (1.1), 1-12
0-2 (%)	241 (62.9)	768 (55.9)	295 (31.8)	83 (7.9)	
Running speed during training km/hr, mean (SD), range #	8.9 (1.8), 5-16	10 (1.7), 5-25	10.8 (1.4), 5-17	11.0 (1.4), 5-21	10.4 (1.7) 5-25
Running experience, years, median (IQR), range #	2 (1-7), 0-45	4 (2-11), 0-48	5 (3-12), 0-51	8 (4-18), 0-56	5 (2-13) 0-56
0-2 year, n (%)	226 (59.0)	551 (40.1)	207 (22.3)	147 (13.9)	
Hard training underground: often (%) #	308 (80.4)	1184 (86.2)	813 (87.8)	969 (91.8)	3298 (87.5)
Long-distance training: often (%) #	306 (79.9)	1241 (90.3)	864 (93.2)	994 (94.2)	3430 (91.0)
Interval training: often (%) #	120 (31.3)	497 (36.2)	417 (45.0)	441 (41.8)	1484 (39.4)
Warming-up before training: often (%)#	206 (53.8)	651 (47.4)	424 (45.7)	417 (39.5)	1711 (45.4)
Stretching before training: often (%)#	194 (50.7)	700 (50.9)	453 (48.9)	423 (40.1)	1783 (47.3)
Cooling down after training: often (%)#	220 (57.5)	666 (48.5)	385 (41.4)	363 (34.4)	1650 (43.8)
Stretching after training: often (%)#	262 (68.4)	918 (66.8)	577 (62.2)	549 (52.0)	2323 (61.6)
Organized running in groups: yes (%)#	114 (29.8)	458 (33.3)	395 (42.6%)	498 (47.2)	1477 (39.2)
Shoe advice: yes (%) #	279 (72.8)	945 (79.4)	806 (86.9)	965 (91.5)	3177 (84.3)

SD: standard deviation; IQR: interquartile range; BMI: body mass index; kg: kilogram; m: meter; km: kilometers; h: hour

*Cumulating numbers do not match because of incidental missings.

Significant differences between groups

Table 2: Characteristic of responders versus non-responders

	Responders T1 **N=2763**	**Non-responders** **N=1005**
Gender, male	1698 (61.5%)	572 (56.9%)*
Age, mean (SD)	43.5 (11.1)	40.8 (11.2)*
BMI, mean (SD)	23.3 (2.4)	23.5 (2.6)*
Running distance*		
5 km	253 (9.2%) [#]	130 (12.9%)[@]
10 km	1000 (36.2%)	374 (37.2%)
Half marathon	713 (25.8%)	214 (21.3%)
Marathon	780 (28.2%)	275 (27.4%)

* means statistical significant difference ($p < 0.05$)

[#] = % runners within -responders

[@] = % runners within -non-responders

Table 3. Multivariate regression models (backward wald*) for running injuries

Variables	5 km (n = 220, 66 injuries)	10 -15 km (n = 818, 224 injuries)	Half marathon (n = 683, 206 injuries)	Marathon (n = 673, 230 injuries)	Total (n = 2369, 709 injuries)
Running distance during the event (categorical)					1.3 (1.2 – 1.5)
Age (continuous, year)	0.97 (0.95 - 0.99)	0.98 (0.97 - 0.99)			0.99 (0.98 - 1)
Previous injury (yes/no)	4.1 (2.2 - 7.6)	3.8 (2.7 - 5.3)	3.3 (2.3 - 4.8)	4.3 (2.9 - 6.1)	3.7 (3.0 - 4.5)
Weekly training distance (continuous, km)	0.95 (0.9 – 0.99)	0.97 (0.95 - 0.99)	0.98 (0.97 - 1)	0.98 (0.97 – 0.99)	0.99 (0.98 – 1)
BMI		1.1 (1.0 – 1.2)			
Weekly training frequency (continuous, nr)		1.3 (0.99 – 1.7)			
Performance measures					
Nagelkerke R square	15.6%	13.4%	9.6%	13.8%	12.1%
Hosmer -Lemeshow	0.89	0.92	0.12	0.85	0.70
Percentage correctly classified	76.7%	72.2%	70%	66.7%	70.2%
AUC (95% CI)	0.71 (0.64-0.79)	0.70 (0.66-0.73)	0.67 (0.62-0.71)	0.68 (0.64-0.72)	0.68 (0.66-0.71)

Data presented as OR (95% CI) unless otherwise specified; OR > 1.00 is a risk factor; OR < 1.00 is a protective factor; CI, confidence interval;

* Exclusion multivariate model p < 0.10;

Table 4 – Nomogram

		Score
Age (per 10 years) [1]	- 2	
Previous injuries	+ 20	
Weekly training (per 10 km)	- 2	
Running distance[2]	+4	
	Total score	

Total score	Probability

Commented [AV4]: How can we calculate this?

[1] The score decreases with 2 points per 10 year (e.g. a 40-year old person receives a score of 4 x -2= -8 points). The same holds for weekly training.

[2] The score increases with 4 point for a running distance of 10-15 km, 8 point for half marathons and 12 point for whole marathons