12-5-2022

Biomechanics associated with tibial stress fracture in runners: A systematic review and meta-analysis

Clare E. Milner

Eric Foch

Joseph M. Gonzales

Drew Peterson

Journal Pre-proof

Biomechanics associated with tibial stress fracture in runners: A systematic review and meta-analysis

Clare E. Milner , Eric Foch , Joseph M Gonzales , Drew Petersen

Please cite this article as: Clare E. Milner , Eric Foch , Joseph M Gonzales , Drew Petersen , Biomechanics associated with tibial stress fracture in runners: A systematic review and meta-analysis, *Journal of Sport and Health Science* (2022), doi: https://doi.org/10.1016/j.jshs.2022.12.002

Review

Biomechanics associated with tibial stress fracture in runners:
A systematic review and meta-analysis

Clare E. Milner [a]*, Eric Foch [b], Joseph M Gonzales [a], Drew Petersen [a]

[a] Department of Physical Therapy and Rehabilitation Sciences, Drexel University, Philadelphia, PA 19102, USA

[b] Department of Health Sciences, Central Washington University, Ellensburg, WA 98926, USA

*Corresponding author: Clare E. Milner

Email milner@drexel.edu

Running title: Running biomechanics and tibial stress fracture

Highlights

- Tibial stress fracture (TSF) is an overuse running injury with a long recovery.
- Many studies refer to biomechanical risk factors for TSF, but only 14 reports have compared biomechanics in runners with TSF to controls.
- Meta-analysis indicated vertical impact peak, vertical active peak, and peak braking force were not statistically significantly different between runners with TSF and controls.
- No conclusive biomechanical risk factors but several variables with moderate or large effects were identified for further investigation.
- Studies may have been underpowered to detect differences. We encourage future studies to include larger samples sizes utilizing multi-center collaborations as appropriate to achieve this.

Graphical Abstract

Biomechanics associated with tibial stress fracture in runners: A systematic review and meta-analysis

Aim: Evaluate biomechanics in runners with tibial stress fracture (TSF) compared to controls by evaluating and synthesizing the peer-reviewed literature

- Overuse running injury with long recovery
- Negatively impacts health and fitness or competition goals
- Many studies refer to biomechanical risk factors for TSF, but most evaluate healthy runners

- Database search identified 359 unique records
- Only 14 included runners with tibial stress fracture
- Meta-analysis found no statistically significant difference in
  i) vertical impact force,
  ii) peak active vertical force or
  iii) peak braking force

- No conclusive biomechanical risk factors, variables for further investigation identified
- Studies may have been underpowered due to small sample sizes
- We encourage future studies with larger sample sizes utilizing multi-center collaborations

Clare E Milner, Eric Foch, Joseph M Gonzales, Drew Petersen, *J Sport Health Sci*, 2022

Abstract

Background: Tibial stress fracture (TSF) is an overuse running injury with a long recovery period. While many running studies refer to biomechanical risk factors for TSF, only a few have compared biomechanics in runners with TSF to controls. The aim of this systematic review and meta-analysis was to evaluate biomechanics in runners with TSF compared to controls.

Methods: Electronic databases PubMed, Web of Science, SPORTDiscus, Scopus, Cochrane, and CINAHL were searched. Risk of bias was assessed and meta-analysis conducted for variables reported in 3 or more studies.

Results: The search retrieved 359 unique records, but only the 14 that compared runners with TSF to controls were included in the review. Most studies were retrospective, 2 were prospective, and most had a small sample size (5–30 per group). Many variables were not significantly different between groups. Meta-analysis of peak impact, active, and braking ground

reaction forces found no significant differences between groups. Individual studies found larger tibial peak anterior tensile stress, peak posterior compressive stress, peak axial acceleration, peak rearfoot eversion and hip adduction in the TSF group.

Conclusion: Meta-analysis indicated that discrete ground reaction force variables were not statistically significantly different in runners with TSF compared to controls. In individual included studies, many biomechanical variables were not statistically significantly different between groups. However, many were reported by only a single study, and sample sizes were small. We encourage additional studies with larger sample sizes of runners with TSF and controls and adequate statistical power to confirm or refute these findings.

Keywords: Bone stress injury; Gait; Kinematics; Kinetics; Tibial acceleration

1. Introduction

Running is a popular form of exercise with many health benefits, but it is also associated with a high rate of overuse injury, ranging from 19% to 80%.[1] Overuse injuries result in time lost from running, which can impact health, well-being, and fitness or competition goals. While many factors both internal and external to the body may contribute to overuse injury,[2] running biomechanics is a readily modified factor and, therefore, a common target for injury prevention efforts.

Tibial stress fracture (TSF) is a running injury caused by repeated mechanical loading leading to bone strain that creates microcracks at a rate that accumulates beyond the bone's capacity for repair and remodeling.[3] It is also a serious injury with a typical recovery period of up to 8 weeks.[4] Furthermore, runners are 5 times more likely to experience a recurrence of stress fracture after the initial injury episode, pointing to an underlying factor that is not resolved during rehabilitation treatments.[5] Thus, efforts to reduce the risk of TSF in runners are needed to break the cycle of recurrent and long-lasting periods of injury. Given the frequency and severity of TSF, running biomechanics have been a target of TSF research.

We have observed that many studies refer to biomechanical risk factors for TSF when interpreting findings on healthy runners, but only a few have compared biomechanics in runners with TSF to controls. Rather, many studies report biomechanics of healthy runners only and do not include a TSF group. Additionally, existing systematic reviews evaluating the literature on running biomechanics and injury have only considered vertical ground reaction force variables.[6,7] Thus, there is a need to systematically review the literature that compares the biomechanics of runners with TSF to controls to determine the strength of current evidence for biomechanical differences and to identify gaps in the literature as well as areas where further research is needed. Therefore, the aim of this systematic review and meta-analysis was to examine biomechanics in runners with TSF compared to controls by evaluating and synthesizing the peer-reviewed literature.

2. Methods

2.1. Literature search

We conducted a systematic review of the published peer-reviewed literature reporting running biomechanics associated with TSF. The review and protocol were not registered but were conducted according to published Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines.[8–10] A research question was developed according to the 3 elements detailed by Booth et al.[11] (who: runners, what: TSF, how: biomechanics): Which biomechanical parameters distinguish runners with TSF from runners without TSF? A search strategy was developed with the assistance of a librarian to define search terms for the study participants (runners), target condition (TSF), and outcome measures of interest (biomechanical variables measured during running). The electronic databases PubMed, Web of Science, SPORTDiscus, Scopus, Cochrane, and CINAHL were searched for published peer-reviewed articles and abstracts from all years up to May 2021. Literature review articles and articles in languages other than English were excluded. The complete search strategy for all databases is detailed in Supplementary Table 1. Additionally, a hand search of the reference lists of review articles identified during the search was conducted.

4

Two reviewers (CM and DP) independently screened items for inclusion in the review in 3 rounds. First, articles that could be excluded based on title were excluded from further review. Second, articles were excluded based on their abstract. Third, the full text of all remaining articles was retrieved, then articles were excluded following review of the full text by both reviewers. The final lists of potential articles for inclusion from each reviewer were compared. Disagreements in article selection were resolved by discussion and joint review of the full text until consensus was reached.

2.2. Data extraction

Study details were extracted into a spreadsheet independently by 2 reviewers (DP and JG). Details extracted included country, year of publication, group definitions and inclusion/exclusion criteria, participant characteristics (weekly mileage, sex, running level, foot strike pattern of participants), sample sizes, experimental location, running velocity, footwear during testing, and primary data analysis. Extracted details for each study were then compared, and discrepancies between reviewers were resolved by discussion and review of the article with the third reviewer (CM).

Biomechanical data for the comparison of TSF and control groups were extracted from all included articles by a single reviewer (JG) and put into a spreadsheet. Group means, standard deviations, and sample sizes were extracted, plus effect size, if reported, and p values for group comparisons. A second reviewer (DP) confirmed the extracted data's fidelity with the original articles. Any discrepancies between reviewers were resolved by discussion and review of the articles with the third reviewer (CM). All quantitative biomechanical variables reported in the articles were extracted. When an outcome variable was reported more than once from the same large research study, only the findings from the report with the largest sample size were included to avoid over-representing the study in this review. Cohen's d effect sizes[12] were calculated when group means and standard deviations were provided in studies that did not report effect size. Data were compiled into tables for presentation of results.

2.3. Risk of bias assessment

The risk of bias in included articles was assessed by 2 reviewers (JG and DP) using 2 tools. Included studies were evaluated according to the Joanna Briggs Institute checklist for analytical cross-sectional studies.[13] The checklist was modified by removing an item about exposure, leaving a total of 7 items for appraisal. Articles were scored 0–7 with 1 point given for each "yes" answer to checklist questions about study methods and statistical analysis. Answers of "no" or "unclear" were given 0 points. Included studies were also evaluated for the quality of study design, reporting of results, and risk of bias using the AXIS tool for cross-sectional studies.[14] Any discrepancies between reviewers were resolved by discussion and review of the article with the third reviewer (CM). An overall risk of bias was determined for each study based on the collective findings of these evaluations.

2.4. Meta-analysis

To be included in the meta-analysis, a variable must have been reported for both TSF and control groups in 3 or more studies. Mean, standard deviation, and group sample size for each variable were entered into the software Review Manager Version 5.41. (RevMan, Copenhagen, Denmark). Separate meta-analyses were performed for each continuous variable. Group mean differences were analyzed via an inverse variance fixed-effect model.[15] This statistical model weights the effect of each study by the inverse of the variance from each study included in the meta-analysis. Group mean differences were considered different from 0 if the overall effect was $p < 0.05$. Ninety-five percent confidence intervals were also computed for the mean differences within each study. The test statistics χ^2 (with corresponding p value) and heterogeneity (I^2) were used to describe the amount of heterogeneity across studies in each meta-analysis.[15] I^2 was considered low (25%), moderate (50%), and high (75%).[15]

6

3. Results

3.1. Study selection

The initial search retrieved 684 records, resulting in 359 unique records when duplicates were removed (Fig. 1). Following evaluation of title and abstract, 337 items were excluded, and 28 full texts were retrieved for assessment, with 14 items (12 research articles and 2 conference abstracts) retained for inclusion in the review.

3.2. Study characteristics

The 14 included publications were published between 1993 and 2020 in Australia, Canada, the UK, and the USA (Table 1). Of these 14, 6 were from the same larger research study reporting runners with a history of TSF.[16–21] Sample sizes ranged from 5–30 participants per group, and *a priori* power analysis for sample size justification was provided in 6 studies (Table 1). All studies compared a group of runners with current, future, or a history of TSF to a healthy control group. There was some variation in how the groups were defined according to study inclusion and exclusion criteria. Occurrence of TSF was confirmed by a medical professional and/or confirmed by imaging in all but a single study. There was also some variation in the definition of the control group. Runners in both groups were healthy and free of injury at the time of data collection in most studies. One study included currently injured runners,[22] and 2 others were prospective.[23,24] Most studies were conducted indoors, looked at overground running, and reflected the traditional gait analysis laboratory setting. One study was conducted on an indoor running track with the runner making contact with the force platform once per lap,[25] and another was conducted on an instrumented treadmill.[22] When reported, running velocity was fixed and ranged from 3.6 m/s to 4.0 m/s. In the treadmill study, running velocity was self-selected and averaged 2.60 m/s in runners with current TSF and 2.65 m/s in controls.[22]

Discrete biomechanical outcome variables were measured and/or calculated in all studies. Direct measurements were made of lower extremity kinematics, tibial acceleration, and ground reaction forces during running. Lower extremity kinetics were calculated via inverse dynamics. Bone

7

stress variables were modeled from 3-dimensional gait analysis combined with bone parameters determined from tibial x-rays. Following traditional gait analysis methods, peak magnitudes for variables of interest or magnitude at defined time points in the stride cycle were extracted from the time series data. Magnitudes were averaged across multiple trials per participant, and group differences analyzed to identify statistically significant differences.

3.3. Risk of bias in included studies

The majority (12/14) of articles scored 7/7 for methodological quality according to the Joanna Briggs Institute checklist, indicating a low risk of bias. Two conference abstracts scored 2/7 and 3/7, respectively, indicating greater risk of bias (Supplementary Table 2). The AXIS appraisal tool mirrored these findings, with the 2 abstracts being of lower methodological quality due to their brevity than the majority of studies, which were high quality (Supplementary Table 3). Thus, except for the 2 abstracts, all included studies were considered to have a low risk of bias.

3.4. Findings of included studies

3.4.1. Meta-analysis results

Due to the greater risk of bias, data reported in either of the 2 abstracts were not included in the meta-analysis. Therefore, meta-analyses were conducted for 3 variables: peak vertical impact force, peak vertical active force, and peak braking force. The results present insufficient evidence to reject the null hypothesis of no difference between groups for any of the ground reaction force variables included in the meta-analysis of runners with previous TSF *vs.* controls ($p > 0.05$). Specifically, the meta-analysis for peak vertical impact force included 122 runners and had a *p*-value of 0.92 with moderate heterogeneity ($I^2 = 57\%$; Fig. 2). The meta-analysis for peak vertical active force, included 170 runners and had a p-value of 0.36 with low heterogeneity ($I^2 = 0\%$; Fig. 3). Lastly, the meta-analysis for peak braking force included 170 runners and had a *p* value of 0.53 with low heterogeneity ($I^2 = 0\%$; Fig. 4.)

3.4.2. Systematic review results

A total of 25 kinematic and kinetic variables were reported, with many having no effects and insufficient evidence to reject the null hypothesis of no significant differences between groups (Table 2). However, some significant differences ($p < 0.05$) were noted as follows: Peak eversion angle was significantly larger with a moderate effect in the TSF group compared to controls.[19] At the hip, peak adduction angle was larger in the TSF group, with a large effect compared to controls.[19] Early-stance sagittal plane knee joint stiffness was significantly larger with a large effect[18] in the TSF group compared to controls. Additionally, tibial rotation range of motion was smaller, with a moderate effect (compared to controls) during barefoot running in recruits who went on to sustain TSF.[24]

A total of 38 different ground reaction force variables were reported, with most having no to small effects and insufficient evidence to reject the null hypothesis of no significant differences between groups (Table 3). Peak vertical impact force was significantly smaller in the TSF group compared to controls in ant abstract[25] but not statistically significantly different in other studies.[17,26,27] Similarly, in the same abstract peak propulsive force was significantly smaller in the TSF group,[25] but it was not statistically significantly different in another study.[26] It should be noted that the abstract[25] did not indicate whether standard deviation or standard error of the mean was reported; thus, effect sizes were not calculated here. Vertical instantaneous loading rate was significantly larger in a small prospective TSF group compared to controls reported in another abstract,[23] but it was not statistically significantly different in other studies.[20,22] Peak adduction free moment was significantly larger in the TSF group and with a large effect compared to controls in one study[16] but not statistically significantly different and with no effect in another.[28] Free moment at peak braking force was significantly larger in the TSF group and with a moderate effect compared to controls.[16]

Tibial bone stress during running was reported in a recent study by Meardon et al.[29] (Supplementary Table 4). Peak anterior tensile stress and peak posterior compressive stress at the distal third of the tibia were both significantly larger in the TSF group and with moderate effects compared to controls.

4. Discussion

The purpose of this systematic review and meta-analysis was to examine biomechanics in runners with TSF compared to controls by evaluating and synthesizing the peer-reviewed literature. Overall, 359 relevant publications were found, but only 14 compared runners with TSF to a control group. Risk of bias was low, except for the 2 abstracts, which lacked methodological detail due to their short length. All 359 publications identified in the search included some combination of the tibia/tibial and stress fracture/bony injury terms as indicated in Table 1. However, the vast majority of those studies did not include a group of runners with TSF, despite appearing in a search specifically including TSF in the terms. While there is clearly a great deal of interest in understanding the biomechanics associated with TSF in runners, few studies have adequately addressed this by including runners with the injury.

4.1 Sample size of included studies

Overall, the sample sizes of these studies were rather small. According to G*Power,[30] for an independent samples *t* test with 80% power and $p < 0.05$, a sample size of 26 participants per group would be needed to detect a significant difference between groups for a large effect size, and 64 per group to detect a significant difference for a moderate effect size. Only 2 studies included more than 26 participants per group,[19,20] and none had more than 30 per group. Thus, 10 of the 14 reports were underpowered to detect large effects, and all were underpowered to detect moderate effects. We acknowledge that increasing the sample size in studies of runners with TSF will take more resources, which may be a practical limitation for many researchers seeking to investigate the biomechanics of TSF in runners. These practical limitations are likely major contributing factors to the small sample sizes in many of the included studies. Differences of 15% or 1 standard deviation were used to determine sample size in several included studies. However, if there are important differences between TSF and control groups that are smaller than this (e.g., moderate effect sizes), they are not likely to be identified as significant differences. Thus, the systematic review findings for variables reported by only 1 or 2 studies should be considered preliminary and suggestive of variables that may be further investigated in the effort to understand biomechanical differences between runners with TSF and controls.

4.2 Variables for further investigation

Variables from the present studies with moderate or larger effect sizes (and *p* values greater than 0.05) may also be considered worthy of further investigation. Thus, dependent variables to be considered include: peak hip adduction angle,[20] peak rearfoot eversion,[19] tibial internal rotation and rearfoot eversion at impact peak,[19] sagittal plane average knee stiffness,[17] vertical impact peak,[17] average braking ground reaction force,[26] absolute peak free moment,[20] frontal and sagittal plane vertical ground reaction force active peaks,[28] angle of the frontal plane vertical ground reaction force vector at active peak,[28] and peak heel pressure.[24] Overall, the current body of literature comparing runners with TSF to controls identified several biomechanical variables that may be larger in the TSF group and, therefore, appropriate for further investigation.

Of the 67 dependent variables identified, only 3 were reported in 3 or more higher quality studies, and so only these 3 were included in the meta-analysis. All of these variables were discrete ground reaction force variables—vertical impact, active peaks, peak braking force—and were not different in runners with TSF compared to controls. This is not surprising as ground reaction force is a response to the acceleration of the center of mass of the whole body. Therefore, observed ground reaction forces cannot be attributed to an individual joint or specific body segment, such as the tibia.[31] It is feasible that many different body segment acceleration configurations could result in the same magnitude of ground reaction force peaks. Bone stress is influenced by muscle forces and joint reaction forces in addition to ground reaction forces.[29] Therefore, discrete ground reaction force variables are likely not sensitive enough to indicate differences between TSF and control runners.

Of the remaining 64 ground reaction force, kinematic, kinetic, and bone stress variables, the majority were reported in only a single study. Most variables reported were not significantly different between groups, with only small effects identified. However, the study modeling tibial bone stress found significant moderate to large effects for higher anterior tensile and posterior compressive tibial stress in the TSF group compared to controls.[29] These findings suggest that the tibia is loading differently during running in those with TSF compared to controls. Approaches that seek to characterize the biomechanics of the tibia during running, including bone stress and strain, may prove fruitful in teasing out conditions that increase the magnitude of tibial loading. This may help to identify targets for intervention to reduce injury risk during

11

running.[29] These tibial measures may be most insightful when included with variables characterizing running biomechanics to help determine which gait patterns reduce tibial load.

There was a pattern of frontal plane kinematic differences with greater peak rearfoot eversion and hip adduction angles[19,20] but smaller tibial rotation range of motion during running in the TSF group compared to controls.[24] Differences in frontal plane alignment and transverse plane motion in runners with TSF may alter the distribution of loading on the tibia from that seen in controls.[19] Since rearfoot eversion is coupled with tibial rotation at the subtalar joint,[32] a smaller range of tibial rotation with greater peak rearfoot eversion may increase torsional loads within the tibia and contribute to increased bone stress. Therefore, these differences in lower extremity frontal plane alignment may alter the distribution of forces on the tibia, possibly contributing to the risk of injury.

4.3 Characteristics of study participants

Per our review criteria, all TSF groups included runners with TSF. Most studies published as original research articles were retrospective and included runners who had a previous TSF from which they had recovered and who were injury-free at the time of testing. However, 1 study did include runners who were currently injured.[22] Furthermore, 2 studies with a small number of runners in the TSF group were prospective.[23,24] As is typical for overuse running injuries, no prospective biomechanical studies with a large sample size of injured runners have been reported. However, given that there is a high likelihood of recurrence of TSF following the initial injury,[5] retrospective studies can provide insight into underlying factors that may be associated with the injury and future injury recurrence. Criteria for inclusion in the control group varied from no history of TSF specifically to no history of any running injury. If there are unique biomechanical features of running associated with TSF, the control group would, at a minimum, need to only exclude runners with a TSF. Given the reported 19% to 79% incidence of injury in runners,[1] excluding runners with any previous running injury from the control group greatly reduces the available pool of participants. Thus, to facilitate the inclusion of larger sample sizes, we recommend that future studies include currently healthy runners with confirmed history of TSF in the injury group and runners with no history of TSF in the control group.

The majority of studies focused on female runners, likely because female runners have a higher incidence of stress fracture.[33] However, 3 studies included only male runners,[24,27,28] 2 included

both men and women,[22,29] and 1 did not report.[25] Since differences in running biomechanics between men and women have been reported for some lower extremity biomechanical variables,[34] it is necessary to confirm that differences reported in female runners also occur in male runners and vice versa. For example, a study conducted in female runners comparing those with TSF to a control group found larger peak free moment in the TSF group,[14] but a study comparing male runners found no difference in the same variable between groups.[28] However, it cannot be determined whether these are gender differences or simply conflicting study findings. Therefore, we recommend that future studies include both male and female runners and power the study so that women and men can be treated as separate subgroups during statistical analysis. Alternatively, if limited resources prohibit this, we suggest focusing on either female or male runners.

All but 2 studies reported laboratory gait analysis during overground running in short trials of 15–30 m. Recent work with precision wearables found significant differences between the magnitude of biomechanical variables collected in the traditional gait analysis laboratory setting and field measures.[35] In particular, tibial acceleration variables were higher when measured during outdoor running compared to running in the laboratory in healthy runners.[35] Thus, biomechanical variables associated with TSF during laboratory gait analysis cannot be assumed to have the same magnitudes when measured during running in the field. Field-based investigations must seek to determine whether the same differences exist between runners with TSF and controls when they are in the outdoor environment.

Some limitations of this review should be noted. The search was restricted to research literature published in English and so does not account for studies published in other languages. Our focus was on running biomechanics, which may be modifiable, and so we did not include bone geometry or other unmodifiable anatomical variables. Since our focus was specifically on TSF, we excluded studies that placed runners with stress fracture at other lower extremity locations within the same group as those with TSF. This reduced the number of included studies but avoided the risk of masking differences attributable to TSF that may be inconsistent with other stress fracture sites, such as the femur or metatarsals. Furthermore, given the potential limiting effects on meta-analysis of a small number of included studies and moderate heterogeneity,[36] the finding for peak vertical impact force should be interpreted with caution.

5. Conclusion

The literature reveals an ongoing interest in identifying approaches for reducing the risk of TSF in runners by examining running biomechanics. However, we found only 14 reports (2 of which were abstracts) that compared aspects of running biomechanics between TSF and control groups, and only 1 that reported tibial stress. Many variables were reported by only a single study. Many variables were not statistically significantly different between the TSF and control groups. Specifically, meta-analysis indicated that the discrete ground reaction force variables vertical impact peak, vertical active peak, and peak braking force were not statistically significantly different in runners with TSF compared to controls. Sample sizes were small, so studies may have been underpowered to detect important differences. We encourage future studies to compare runners with previous, current, or prospective TSF to controls with no history of TSF and to use sample sizes of at least 26 per group to detect group differences with large effects and at least 64 to detect moderate effects. This may require multi-center studies to ensure sufficient statistical power. While prospective studies are the gold standard, we acknowledge that the resources required for these studies are substantial. Thus, we also encourage studies comparing runners with a history of TSF or current TSF to controls to identify biomechanics associated with TSF.

Authors' contributions

CM conceived the idea and developed the design of the review, reviewed the literature, wrote the manuscript, and checked data fidelity; JG and DP also reviewed the literature, contributed to table and figure preparation and data extraction, checked data fidelity, provided suggestions and revisions to the original draft; EF performed all meta-analyses and associated figure preparation

14

and provided suggestions and revisions to the original draft. All authors have read and approved the final version of the manuscript and agree with the order of presentation of the authors.

Competing interests

The authors declare that they have no competing interests.

References

1. van Gent RN, Siem D, van Middelkoop M, van Os AG, Bierma-Zeinstra SM, Koes BW. Incidence and determinants of lower extremity running injuries in long distance runners: A systematic review. *Br J Sports Med* 2007;**41**:469-80.

2. Bennell K, Matheson G, Meeuwisse W, Brukner P. Risk factors for stress fractures. *Sports Med* 1999;**28**:91-122.

3. Warden SJ, Davis IS, Fredericson M. Management and prevention of bone stress injuries in long-distance runners. *J Orthop Sports Phys Ther* 2014;**44**:749-65.

4. Bennell K, Brukner P. Preventing and managing stress fractures in athletes. *Phys Ther Sport* 2005;**6**:171-80.

5. Wright AA, Taylor JB, Ford KR, Siska L, Smoliga JM. Risk factors associated with lower extremity stress fractures in runners: A systematic review with meta-analysis. *Br J Sports Med* 2015;**49**:1517-23.

6. Zadpoor AA, Nikooyan AA. The relationship between lower-extremity stress fractures and the ground reaction force: A systematic review. *Clin Biomech (Bristol, Avon)* 2011;**26**:23-8.

7. van der Worp H, Vrielink JW, Bredeweg SW. Do runners who suffer injuries have higher vertical ground reaction forces than those who remain injury-free? A systematic review and meta-analysis. *Br J Sports Med* 2016;**50**:450-7.

8. Ardern CL, Büttner F, Andrade R, et al. Implementing the 27 PRISMA 2020 statement items for systematic reviews in the sport and exercise medicine, musculoskeletal rehabilitation and sports

science fields: the PERSiST (implementing PRISMA in exercise, rehabilitation, sport medicine and sports science) guidance. *Br J Sports Med* 2022;**56**:175-95.

9. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;**372**:n160. doi:10.1136/bmj.n160.

10. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71. doi:10.1136/bmj.n71.

11. Booth A, Papaioannou D, Sutton A. *Systematic approaches to a successful literature review.* London: Sage; 2012

12. Cohen J. A power primer. *Psychol Bull* 1992;**112**:155-9.

13. Moola S, Munn Z, Tufanaru C, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, editors. *JBI Manual for Evidence Synthesis.* JBI, 2020. Available at: https://synthesismanual.jbi.global. doi:10.46658/JBIMES-20-08.

14. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open* 2016;**6**:e011458. doi:10.1136/bmjopen-2016-011458.

15. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557-60.

16. Milner CE, Davis IS, Hamill J. Free moment as a predictor of tibial stress fracture in distance runners. *J Biomech* 2006;**39**:2819-25.

17. Milner CE, Ferber R, Pollard CD, Hamill J, Davis IS. Biomechanical factors associated with tibial stress fracture in female runners. *Med Sci Sports Exerc* 2006;**38**:323-8.

18. Milner CE, Hamill J, Davis I. Are knee mechanics during early stance related to tibial stress fracture in runners? *Clin Biomech (Bristol, Avon)* 2007;**22**:697-703.

19. Milner CE, Hamill J, Davis IS. Distinct hip and rearfoot kinematics in female runners with a history of tibial stress fracture. *J Orthop Sports Phys Ther* 2010;**40**:59-66.

20. Pohl MB, Mullineaux DR, Milner CE, Hamill J, Davis IS. Biomechanical predictors of retrospective tibial stress fractures in runners. *J Biomech* 2008;**41**:1160-5.

21. Zifchock RA, Davis I, Hamill J. Kinetic asymmetry in female runners with and without retrospective tibial stress fractures. *J Biomech* 2006;**39**:2792-7.

22. Johnson CD, Tenforde AS, Outerleys J, Reilly J, Davis IS. Impact-related ground reaction forces are more strongly associated with some running injuries than others. *Am J Sports Med* 2020;**48**:3072-80.

23. Davis I, Milner C, Hamill J. Does increased loading during running lead to tibial stress sractures? A prospective study. *Med Sci Sports Exerc* 2004;**36**:S58. doi:10.1249/00005768-200405001-00271.

24. Nunns M, House C, Rice H, et al. Four biomechanical and anthropometric measures predict tibial stress fracture: A prospective study of 1065 Royal Marines. *Br J Sports Med* 2016;**50**:1206-10.

25. Grimston SK, Nigg BM, Fisher V, Ajemian SV. External loads throughout a 45 minute run in stress fracture and non-stress fracture runners. *Proceedings of the International Society of Biomechanics XIV Congress.* Paris, France; July 1993.

26. Bennell K, Crossley K, Jayarajan J, et al. Ground reaction forces and bone parameters in females with tibial stress fracture. *Med Sci Sports Exerc* 2004;**36**:397-404.

27. Crossley K, Bennell KL, Wrigley T, Oakes BW. Ground reaction forces, bone characteristics, and tibial stress fracture in male runners. *Med Sci Sports Exerc* 1999;**31**:1088-93.

28. Creaby MW, Dixon SJ. External frontal plane loads may be associated with tibial stress fracture. *Med Sci Sports Exerc* 2008;**40**:1669-74.

29. Meardon SA, Willson JD, Gries SR, Kernozek TW, Derrick TR. Bone stress in runners with tibial stress fracture. *Clin Biomech* 2015;**30**:895-902.

30. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;**39**:175-91.

31. Winter D. *Biomechanics and Motor Control of Human Movement.* Hoboken, NJ: John Wiley & Sons, Inc.; 2009.p.121.

32. Inman VT. The influence of the foot-ankle complex on the proximal skeletal structures. *Artif Limbs* 1969;**13**:59-65.

33. Wentz L, Liu PY, Haymes E, Ilich JZ. Females have a greater incidence of stress fractures than males in both military and athletic populations: A systemic review. *Mil Med* 2011;**176**:420-30.

34. Ferber R, Davis IM, Williams 3rd DS. Gender differences in lower extremity mechanics during running. *Clin Biomech (Bristol, Avon)* 2003;**18**:350-7.

35.   Milner CE, Hawkins JL, Aubol KG. Tibial acceleration during running is higher in field testing than indoor testing. *Med Sci Sports Exerc* 2020;**52**:1361-6.

36.   Seide SE, Röver C, Friede T. Likelihood-based random-effects meta-analysis with few studies: Empirical and simulation studies. *BMC Med Res Methodol* 2019;**19**:16. doi:10.1186/s12874-018-0618-3.

18

Table 1. Description of included studies.

Study and year	Participants and sample size	*A priori* power analysis?	Weekly mileage (mean ± SD)	Type of runner and foot strike pattern (if reported)	Experimental location	Velocity (m/s) and footwear (if reported)
Grimston et al., (1993)[25]	5 retrospective TSF (definition NR), 5 control (no history of stress fracture)	No	NR	NR	Indoor track, *early* and *late* stages of 45-min run.	NR
Crossley et al., (1999)[27]	23 currently healthy, retrospective TSF (physician diagnosis, confirmed by imaging), 23 currently healthy control (no history of stress fracture)	No, *post hoc* 95% power to detect difference of 10% between groups	NR	Male rearfoot strike running athletes	30 m, indoor, over-ground	4.0 ± 10%; own running shoes
Bennell et al., (2004)[26]	13 currently healthy, retrospective TSF (physician diagnosis, confirmed by imaging), 22 currently healthy control (no history of stress fracture)	No, *post hoc* sample size sufficient to detect 10% difference between groups	TSF: 33.7 ± 20.1; control: 30.3 ± 21.6	Female running athletes	30 meters, indoor, over-ground	4.0 ± 0.4; own running shoes
Davis et al., (2004)[23]	5 prospective TSF (documented TSF or tibial stress reaction), 5 control (definition NR)	No	NR	Competitive women runners	25 m, indoor, over-ground in laboratory	3.8
Milner et	20 currently healthy,	Yes, 20 per	TSF: 28.6	Recreational and	23 m, indoor, over-	3.7 ± 5%;

al., (2006)[17]	retrospective TSF (medical professional diagnosis, confirmed by imaging), 20 currently healthy control (no previous lower extremity bony injury)	group to detect 15% difference, alpha 0.05, 80% power using preliminary data	± 6.8; control: 29.2 ± 9.9	competitive rearfoot strike female runners	ground in laboratory	neutral running shoes
Milner et al., (2006)[16]	25 currently healthy, retrospective TSF (medical professional diagnosis, confirmed by imaging), 25 currently healthy control (no previous lower extremity fractures)	Yes, 24 per group for effect size 0.48, α 0.05, 80% power using cited preliminary data	TSF: 28.6 ± 9.3; control: 28.6 ± 11.8	Recreational and competitive rearfoot strike female runners	23 m, indoor, over-ground	3.7 ± 5%; neutral running shoes
Zifchock et al., (2006)[21]	24 currently healthy, retrospective TSF (history of 1 or more unilateral TSF, confirmed by imaging), 25 currently healthy control (never sustained a running injury)	Yes, 24 for α 0.05, 80% power using data from literature	TSF: 26.8 ± 10.8; Control: 28.8 ± 11.8	Female rearfoot strike runners	25 m, indoor, over-ground	3.7 ± 5%; neutral running shoes
Milner et al., (2007)[18]	23 currently healthy, retrospective TSF (medical professional diagnosis, confirmed by imaging), 23 currently healthy control (no	Yes, 19 per group to detect 15% difference, α 0.05, 80% power using	TSF: 29.2 ± 8.7; control: 28.6 ± 9.3	Recreational rearfoot strike female runners	23 m, indoor, over-ground in laboratory	3.7 ± 5%; neutral running shoes

	previous lower extremity bony injury)	preliminary data				
Creaby et al., (2008)[28]	10 currently healthy, retrospective TSF (confirmed by Principal Medical Officer using an imaging algorithm), 20 currently healthy control (no history of lower limb injury sustained during training)	No	NR	Male rearfoot strike military recruits	20 m, indoor, over-ground	3.6 ± 5%; above ankle combat assault boots
Pohl et al., (2008)[20]	30 currently healthy, retrospective TSF (confirmed by a medical doctor), 30 currently healthy control (no previous lower extremity bony injury)	No	TSF: 25.5 ± 6.8; control: 24.2 ± 8.7	Female rearfoot strike runners	25 m, indoor, over-ground	3.7 ± 5%; neutral running shoes
Milner et al., (2010)[19]	29 currently healthy, retrospective TSF (medical professional diagnosis, confirmed by imaging), 29 currently healthy control (no previous lower extremity bony injury)	Yes, 29 per group to detect difference of 1 standard deviation, α 0.05, 80% power using preliminary data	TSF: 28.6 ± 13.0; Control: 26.7 ± 7.5	Female rearfoot strike running athletes	Indoor, over-ground in laboratory	3.7 ± 5%; neutral running shoes
Meardon et al.,	23 currently healthy, retrospective TSF (physician	Yes, 23 per group to detect	TSF: 18.2 ± 10.7;	Male or female runners	25 m, indoor, over-ground	3.7; neutral running

21

(2015)[29]	diagnosis, confirmed by imaging), 23 currently healthy control (no history of stress fracture)	15% difference, alpha 0.05, 80% power using data from literature	control: 17.8 ± 10.6			shoes
Nunns et al., (2016)[24]	10 prospective TSF (confirmed by imaging), 150 currently healthy control (recruits who completed training without injury)	No	NR	Military recruits	15 m, indoor, over-ground	3.6 ± 5%; barefoot
Johnson et al., (2020)[22]	23 currently injured TSF (tibial bone stress injury, confirmed by imaging, pain ≤ 2/10), 65 currently healthy control (injury-free for previous 3 months)	No	NR	Male or female rearfoot strike runners	Instrumented treadmill	Self-selected velocity. TSF (2.65 ± 0.24), Controls (2.60 ± 0.22); lab shoes matched to habitual footwear

Abbreviations: NR = not reported; TSF = tibial stress fracture.

Table 2. Kinematic and kinetic variables in tibial stress fracture (TSF) and control groups.

Variable	Study	TSF (mean ± SD)	Control mean (SD)	Effect size	p
Peak rearfoot eversion angle (°)	Milner et al.[19]	11.7 ± 4.2	9.0 ± 3.9	0.66	0.015
Impact peak rearfoot eversion angle (°)	Milner et al.[19]	5.5 ± 4.0	3.0 ± 4.3	0.61	>0.05
Knee flexion at footstrike (°)	Milner et al.[18]	13.7 ± 6.0	11.9 ± 6.5	0.28	0.348
Knee flexion excursion (°)	Milner et al.[17]	33.1 ± 5.0	34.8 ± 5.2	0.34	0.147
Early stance knee flexion excursion (°)	Milner et al.[18]	14.4 ± 4.0	16.0 ± 5.3	0.36	0.252
Peak knee adduction angle (°)	Milner et al.[19]	1.4 ± 4.0	2.2 ± 5.2	0.17	0.505
Impact peak knee adduction angle (°)	Milner et al.[19]	−1.6 ± 3.6	−0.2 ± 5.1	0.33	>0.05
Peak knee internal rotation angle (°)	Milner et al.[19]	3.9 ± 5.2	3.1 ± 6.9	0.13	0.633
Impact peak knee internal rotation angle (°)	Milner et al.[19]	−2.0 ± 6.3	−4.4 ± 6.4	0.37	>0.05
Shank angle at footstrike (°)	Milner et al.[18]	12.8 ± 3.4	14.1 ± 3.3	0.40	0.181

Peak tibial internal rotation angle (°)	Milner et al.[19]	−9.4 ± 5.8	−6.7 ± 5.6	0.47	0.080
Impact peak tibial internal rotation angle (°)	Milner et al.[19]	−7.1 ± 4.6	−4.2 ± 6.4	0.53	>0.05
Tibial rotation range of motion (°)	Nunns et al.[24]	6.4 ± 4.3	10.3 ± 6.0	−0.66*	0.05
Peak hip adduction angle (°)	Milner et al.[19]	11.6 ± 5.0	8.1 ± 3.7	0.80	0.004
Impact peak hip adduction angle (°)	Milner et al.[19]	6.3 ± 6.6	4.7 ± 4.9	0.29	>0.05
Peak hip internal rotation angle (°)	Milner et al.[19]	6.6 ± 5.0	8.5 ± 6.1	0.33	0.222
Impact peak hip internal rotation angle (°)	Milner et al.[19]	2.4 ± 6.1	3.7 ± 6.9	0.20	>0.05
Peak axial tibial acceleration (g)	Pohl et al.[20]	6.5 ± 3.4	5.5 ± 2.5	0.3	NR
Peak positive tibial acceleration symmetry index (%)	Zifchock et al.[21]	29.0 ± 23.3	31.7 ± 24.1	−0.11*	0.70
Sagittal plane average ankle joint stiffness (x10^{-2})[a]	Milner et al.[17]	4.31 ± 0.59	4.59 ± 0.61	−0.46	<0.05

24

Sagittal plane average knee joint stiffness $(x10^{-2})^a$	Milner et al.[17]	4.88 ± 0.88	4.46 ± 0.68	0.54	0.054
Early stance sagittal plane average knee joint stiffness $(x10^{-2})^a$	Milner et al.[18]	4.4 ± 2.1	3.0 ±1.5	0.79	0.015
Lower extremity stiffness (kN/m)	Davis et al.[23]	9.21 (NR)	9.63 (NR)	NR	0.30
Vertical stiffness at initial loading (kN/m)	Johnson et al.[22]	77.59 ± 22.48	68.32 ± 18.91	0.47*	>0.05
Peak heel pressure (N/cm^2)	Nunns et al.[24]	20.81 ± 6.83	17.69 ± 4.75	0.64*	0.06

[a] Joint stiffness is change in joint moment (Nm/(mass in kg*height in m) divided by change in joint angle (º). *: calculated effect size. Note that when an outcome variable was reported more than once from the same study, only the report with the largest sample size is included here.

Abbreviations: kN = kilonewton; Nm = Newton-meter; NR = not reported; TSF = tibial stress fracture.

Table 3. Ground reaction force variables in tibial stress fracture (TSF) and control groups.

Variable	Study	TSF (mean ± SD)	Control (mean ± SD)	Effect size	p value
Peak vertical impact force (BW)	Grimston et al.[25]	1.84	2.24	–	Significant
	Crossley et al.[27]	1.890 ± 0.387	1.970 ± 0.337	–0.22*	>0.05
	Bennell et al.[26]	1.944 ± 0.295	2.080 ± 0.381	–0.39*	0.32
	Milner et al.[17]	1.84 ± 0.21	1.70 ± 0.32	0.51	0.057
Time to peak vertical impact force (s)	Crossley et al.[27]	0.031 ± 0.005	0.031 ± 0.005	0.00*	>0.05
	Bennell et al.[26]	0.136 ± 0.016	0.132 ± 0.023	0.19*	0.65
Peak vertical active force (BW)	Grimston et al.[25]	2.48	2.68	–	Significant
	Crossley et al.[27]	2.843 ± 0.235	2.856 ± 0.189	–0.06*	>0.05
	Bennell et al.[26]	2.747 ± 0.216	2.786 ± 0.247	–0.16*	0.47
	Davis et al.[23]	2.55 (NR)	2.63 (NR)	NR	0.15
	Johnson et al.[22]	2.24 ± 0.22	2.28 ± 0.22	–0.18*	NR
Time to peak vertical active force (s)	Crossley et al.[27]	0.099 ± 0.009	0.097 ± 0.011	0.20*	>0.05
	Bennell et al.[26]	0.452 ± 0.028	0.451 ± 0.047	0.02*	0.94
Average vertical force (BW)	Bennell et al.[26]	1.654 ± 0.138	1.696 ± 0.130	–0.32*	0.37
Peak braking force (BW)	Grimston et al.[25]	0.28	0.35	–	Significant
	Crossley et al.[27]	–0.496 ± 0.056	–0.492 ± 0.104	–0.05*	>0.05
	Bennell et al.[26]	–0.497± 0.080	–0.515 ± 0.088	0.21*	0.54
	Johnson et al.[22]	0.25 ± 0.06	0.23 ± 0.12	0.19	NR

26

Time to peak braking force (s)	Crossley et al.[27]	0.050 ± 0.008	0.051 ± 0.013	−0.09*	>0.05
	Bennell et al.[26]	0.211 ± 0.063	0.207 ± 0.053	0.07*	0.54
Average braking force (BW)	Bennell et al.[26]	−0.232 ± 0.031	−0.249 ± 0.033	0.53*	0.13
Peak propulsive force (BW)	Grimston et al.[25]	0.37	0.47	–	Significant
	Bennell et al.[26]	0.369 ± 0.064	0.380 ± 0.039	−0.22*	0.55
Time to peak propulsive force (s)	Bennell et al.[26]	0.751 ± 0.016	0.757 ± 0.016	−0.38*	0.35
Average propulsive force (BW)	Bennell et al.[26]	0.213 ± 0.033	0.220 ± 0.021	−0.27*	0.42
Peak medial force (BW)	Johnson et al.[22]	0.083 ± 0.023	0.093 ± 0.068	−0.17	NR
Peak lateral force (BW)	Johnson et al.[22]	0.071 ± 0.037	0.065 ± 0.041	0.15*	NR
Sagittal plane impact peak force (BW)	Creaby et al.[28]	1.91 ± 0.22	1.81 ± 0.26	0.38	0.17
Angle of sagittal plane impact peak vector (°)	Creaby et al.[28]	−5.65 ± 3.23	−5.84 ± 4.67	0.05	OP
Sagittal plane active peak force (BW)	Creaby et al.[28]	2.49 ± 0.18	2.67 ± 0.21	−0.85	OP
Angle of sagittal	Creaby et al.[28]	−3.25 ± 1.63	−2.47 ± 1.69	−0.46	0.12

27

plane active peak vector (°)					
Frontal plane impact peak force (BW)	Creaby et al.[28]	1.9 ± 0.22	1.8 ± 0.26	0.39	0.16
Angle of frontal plane impact peak vector (°)	Creaby et al.[28]	−1.30 ± 3.83	−1.40 ± 3.40	0.03	OP
Frontal plane active peak force (BW)	Creaby et al.[28]	2.49 ± 0.19	2.67 ± 0.20	−0.87	OP
Angle of frontal plane active peak vector (°)	Creaby et al.[28]	−2.34 ± 1.57	−1.39 ± 1.36	−0.64	0.05
Vertical instantaneous loading rate (BW/s)	Davis et al.[23]	112.88 (NR)	81.03 (NR)	NR	0.04
	Pohl et al.[20]	88.2 ± 24.7	83.8 ± 23.2	0.2	NR
	Johnson et al.[22]	70.78 ± 21.55	63.50 ± 20.52	0.35*	>0.05
Vertical average loading rate (BW/s)	Davis et al.[23]	88.20 (NR)	62.91 (NR)	NR	0.06
	Pohl et al.[20]	74.2 ± 23.5	66.0 ± 22.4	0.4	NR
	Johnson et al.[22]	61.18 ± 19.60	54.37 ± 18.25	0.37*	NR
Braking instantaneous loading rate (BW/s)	Milner et al.[17]	20.35 ± 6.17	19.29 ± 4.70	0.19	0.272
	Johnson et al.[22]	8.66 ± 2.86	9.02 ± 5.73	−0.07*	NR
Braking average	Milner et al.[17]	8.54 ± 3.10	8.37 ± 2.25	0.07	0.420

28

loading rate (BW/s)					
Medial instantaneous loading rate (BW/s)	Johnson et al.[22]	8.85 ± 5.47	8.30 ± 4.54	0.11*	NR
Lateral instantaneous loading rate (BW/s)	Johnson et al.[22]	7.82 ± 4.12	7.94 ± 7.45	−0.02*	NR
Peak vertical impact force symmetry index (%)	Zifchock et al.[21]	8.8 ± 13.6	12.6 ± 10.1	−0.32*	0.27
Peak vertical active force symmetry index (%)	Zifchock et al.[21]	2.6 ± 1.8	3.1 ± 2.5	−0.23*	0.42
Peak braking force symmetry index (%)	Zifchock et al.[21]	10.8 ± 12.8	11.4 ± 8.8	−0.05*	0.85
Peak lateral force symmetry index (%)	Zifchock et al.[21]	36.3 ± 42.9	49.8 ± 38.3	−0.33*	0.25
Peak medial force symmetry index (%)	Zifchock et al.[21]	44.8 ± 44.8	37.5 ± 28.0	0.20*	0.50
Peak instantaneous vertical loading rate symmetry index (%)	Zifchock et al.[21]	12.6 ± 9.4	15.0 ± 10.4	−0.24*	0.40
Vertical average loading rate symmetry index (%)	Zifchock et al.[21]	16.5 ± 11.7	23.3 ± 17.4	−0.46*	0.11

29

Peak adduction free moment ($\times 10^{-3}$)	Milner et al.[17]	7.7 ± 4.7	4.7 ± 2.5	0.80	0.004
	Creaby et al.[28]	6.2 ± 2.4	5.7 ± 3.1	0.14	0.35
Free moment at peak braking force ($\times 10^{-3}$)	Milner et al.[17]	4.6 ± 5.7	1.6 ± 3.7	0.62	0.017
Absolute peak free moment ($\times 10^{-3}$)	Pohl et al.[20]	9.1 ± 4.2	6.1 ± 2.5	0.9	NR
	Creaby et al.[28]	9.5 ± 2.1	9.3 ± 3.2	0.01	0.4
Net angular impulse (s, $\times 10^{-4}$)	Milner et al.[16]	4.5 ± 9.9	1.6 ± 5.5	0.36	0.105

* Calculated effect size. Significant: *p* value not provided. Note that when an outcome variable was reported more than once from the same study, only the report with the largest sample size is included here.

Abbreviations: BW = body weight; s = seconds; NR = not reported; OP = opposite of hypothesized direction; TSF = tibial stress fracture; "–" = effect size not calculated due to data uncertainty.
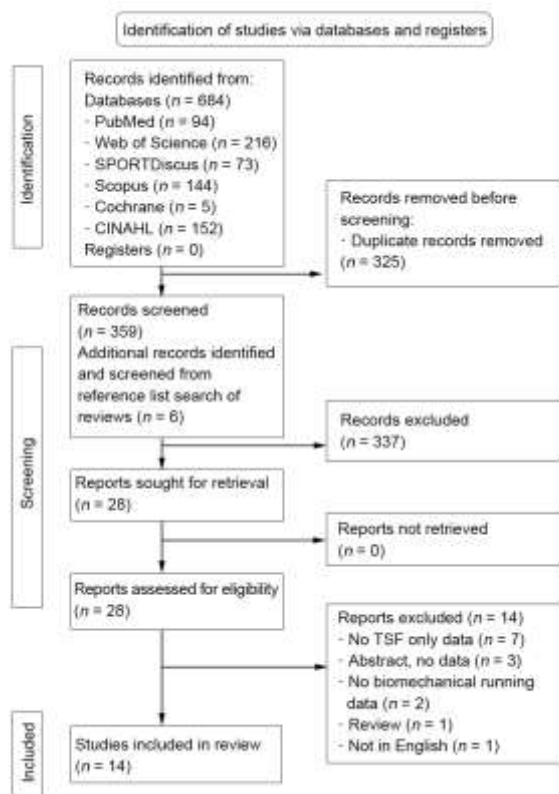
Fig. 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020[10] flow diagram.

Study or subgroup	TSF			Control			Weight	Mean difference	Mean difference (body weights)
	mean	SD	Total	mean	SD	Total		IV, fixed, 95%CI	IV, fixed, 95%CI
Bennell et al.[23]	1.94	0.3	13	2.08	0.38	23	25.4%	-0.14 (-0.37 to 0.09)	
Crossley et al.[24]	1.89	0.39	23	1.97	0.34	23	28.8%	-0.08 (-0.29 to 0.13)	
Milner et al.[14]	1.84	0.21	20	1.7	0.32	20	45.8%	0.14 (-0.03 to 0.31)	
Total (95%CI)			56			66	100.0%	0.01 (-0.11 to 0.12)	

Heterogeneity: $x^2 = 4.70$, df = 2 ($p = 0.10$); $I^2 = 57\%$
Test for overall effect: Z = 0.09 ($p = 0.92$)
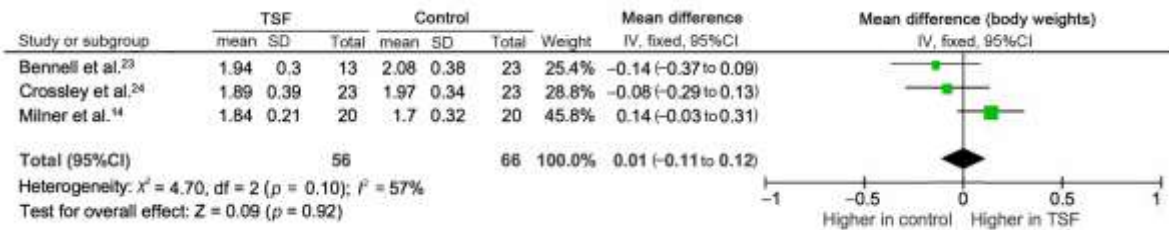
Higher in control   Higher in TSF

Fig. 2. Forest plot of vertical impact peak during running showing no difference between groups. Normalized to body weight. 95%CI = 95% confidence interval; IV = inverse variance; TSF = tibial stress fracture.

Study or subgroup	TSF			Control			Weight	Mean difference	Mean difference (body weights)
	Mean	SD	Total	Mean	SD	Total		IV, fixed, 95%CI	IV, fixed, 95%CI
Bennell et al.[23]	2.75	0.22	13	2.79	0.25	23	20.7%	-0.04 (-0.20 to 0.12)	
Crossley et al.[24]	2.84	0.24	23	2.86	0.19	23	32.7%	-0.02 (-0.15 to 0.11)	
Johnson et al.[19]	2.24	0.22	23	2.28	0.22	65	46.7%	-0.04 (-0.14 to 0.06)	
Total (95%CI)			59			111	100.0%	-0.03 (-0.10 to 0.04)	

Heterogeneity: $x^2 = 0.07$, df = 2 ($p = 0.97$); $I^2 = 0\%$
Test for overall effect: Z = 0.92 ($p = 0.36$)
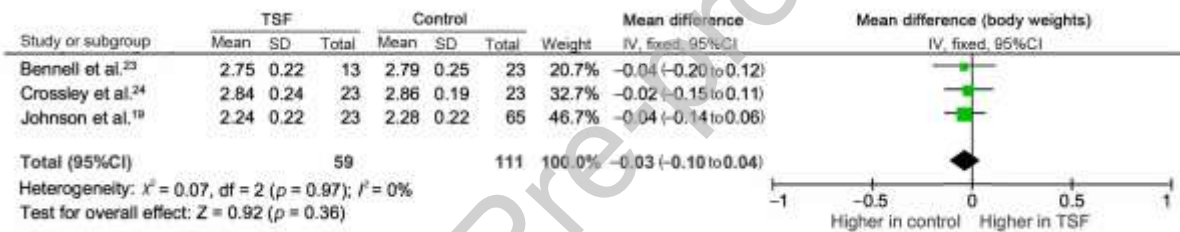
Higher in control   Higher in TSF

Fig. 3. Forest plot of peak vertical active force during running showing no difference between groups. Normalized to body weight. 95%CI = 95% confidence interval; IV = inverse variance; TSF = tibial stress fracture.

32

Study or subgroup	TSF			Control			Weight	Mean difference IV, fixed, 95%CI	Mean difference (body weights) IV, fixed, 95%CI
	mean	SD	Total	mean	SD	Total			
Bennell et al.[23]	0.5	0.08	13	0.52	0.09	23	21.5%	−0.02 (−0.08 to 0.04)	
Crossley et al.[24]	0.5	0.06	23	0.49	0.1	23	30.6%	0.01 (−0.04 to 0.06)	
Johnson et al.[19]	0.25	0.06	23	0.23	0.12	65	47.9%	0.02 (−0.02 to 0.06)	
Total (95%CI)			59			111	100.0%	0.01 (−0.02 to 0.03)	

Heterogeneity: $\chi^2 = 1.32$, df = 2 ($p = 0.52$); $I^2 = 0\%$
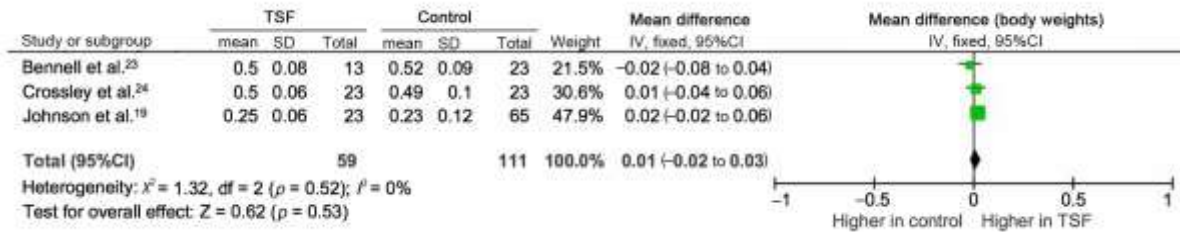Test for overall effect: $Z = 0.62$ ($p = 0.53$)

Fig. 4. Forest plot of peak braking force during running showing no difference between groups. Normalized to body weight. 95%CI = 95% confidence interval; IV = inverse variance; TSF = tibial stress fracture.

33