The effect of surface compliance on overground running biomechanics. A systematic review and meta-analysis

Cameron Mitchell, Sarah McDonnell, Karina Oganezova, David Mockler & Neil Fleming

To cite this article: Cameron Mitchell, Sarah McDonnell, Karina Oganezova, David Mockler & Neil Fleming (2023): The effect of surface compliance on overground running biomechanics. A systematic review and meta-analysis, Sports Biomechanics, DOI: 10.1080/14763141.2023.2236058

To link to this article: https://doi.org/10.1080/14763141.2023.2236058

Published online: 21 Jul 2023.
The effect of surface compliance on overground running biomechanics. A systematic review and meta-analysis

Cameron Mitchell, Sarah McDonnell, Karina Oganezova, David Mockler and Neil Fleming

© Department of Anatomy, School of Medicine, Trinity College, The University of Dublin, Dublin, Ireland; John Stearne Medical Library, Trinity Centre for Health Sciences, School of Medicine, St. James’s Hospital, Dublin, Ireland

ABSTRACT
The surface upon which running is performed has been suggested as a potential cause of many running-related injuries. It remains unclear, however, what effect surface compliance has on running biomechanics. This study aimed to investigate the effect of surface compliance on overground running biomechanics through a systematic review and meta-analysis. Using the PRISMA Protocols Statement, a search was conducted in three electronic databases (CINAHL, EMBASE, EBSCO) using the following anchoring terms: running, overground surface, biomechanics, kinematics, tibial acceleration, pressure and force. Following de-duplication, title/abstract screening and full-text review, 25 articles (n = 492) were identified which met all inclusion criteria, 22 (n = 392) of which were subsequently included in quantitative synthesis. Random effects analysis found that peak tibial acceleration was significantly lower when running on softer surfaces (P = 0.01, Z = 2.51; SMD = −0.8; 95% CI = [−1.42 to −0.18]). However, peak vertical ground reaction force, loading rate and ground contact time were not significantly different when comparing hard and soft surfaces. Since peak tibial acceleration has been associated with an increased risk of tibial stress injuries, the results of this meta-analysis suggest that running on softer surfaces to reduce impact stress on the tibia is probably justified to lower the risk of running-related stress injuries.

ARTICLE HISTORY
Received 27 December 2022
Accepted 7 July 2023

KEYWORDS
Running; surface compliance; tibial acceleration

Introduction
Running is one of the world’s most popular forms of exercise (Scheerder et al., 2015), which over the past three decades has seen a profound rise in participation in the developed world (van Dyck et al., 2017). This is in part due to its accessibility and the limited equipment required. Running is also one of the most effective forms of physical activity to not only improve cardiovascular health but also increase caloric expenditure (Kravitz et al., 1997; Thomas et al., 1995). Despite the significant health benefits, running has one of the highest injury incidences of any form of exercise. The reported incidence of running-related injuries (RRI) ranges from 19% to 79% (van Gent et al., 2007) or 2.5 to 33.0 injuries per 1000 h, with novice and recreational runners having a significantly higher risk compared to...
more experienced or elite runners (Kepler et al., 2018; Videbæk et al., 2015). In terms of site of injury, Kakouris et al. (2021) reported that 51.3% RRI occur at or below the knee.

The elevated risk of RRI in novice and recreational runners is still poorly understood. It is likely that multiple factors are contributing, including body mass index (BMI), running experience, running kinematics, training characteristics and footwear (Knapik et al., 2016; Malisoux et al., 2015; Van Middelkoop et al., 2008; van Poppel et al., 2014, 2016, 2018). Running surfaces have also been implicated as a probable cause of RRI in clinical reports and research studies (Boey et al., 2017; Dixon et al., 2000; Tessutti et al., 2012; Tillman et al., 2002). Data from epidemiological studies suggest a relationship exists between surface and the aetiology of injuries (Nigg & Yeaton, 1987). When exploring this relationship, four biomechanical running parameters have frequently been used: peak acceleration of the tibia, ground reaction forces (GRF), loading rates and plantar pressure, all of which may be influenced by the compliance of the surface underfoot (Boey et al., 2017; Dixon et al., 2000; Fu et al., 2015; Tessutti et al., 2012).

Higher vertical ground reaction forces (vGRF), loading rate and tibial acceleration have been associated with a higher risk of RRI such as medial tibial stress syndrome (MTSS), iliotibial band syndrome, patellofemoral syndrome, metatarsal and tibial stress injuries (Boey et al., 2017; Gerlach et al., 2005; van Mechelen, 1992). While a plantar distribution which is more laterally directed at toe off and less pronation on heel strike is known to be a risk factor for overuse RRI (Ghani Zadeh Hesar et al., 2009).

Despite a myriad of biomechanical risk factors being examined, a recent systematic review highlighted that a previous RRI remains the strongest risk factor for a subsequent injury (van Poppel et al., 2021). Considering the considerable risk of a repeat RRI, rehabilitation must be managed appropriately to mitigate reinjury risk. A controlled and progressive reintroduction of running—particularly on soft surfaces in the early phase of rehabilitation—is often advocated (Liem et al., 2013). The rationale assumes that the compliance of the ground underfoot is directly influencing impact loading, which is a known factor for several RRI (Eckard et al., 2018). Therefore, increasing surface compliance may reduce risk of reinjury in the early rehabilitation phase, via an assumption that impact forces may be more effectively attenuated in the early stance phase. However, early biomechanical comparisons of running surface compliance found no significant differences in the resultant impact forces (Feehery, 1986; Nigg & Yeaton, 1987).

Despite several more recent studies examining the interactions between surface compliance and running mechanics (Hong et al., 2012; Shi et al., 2019), the overall effects remain unclear. Some studies have reported that running on hard concrete increased plantar pressure (Wang et al., 2012) and tibial loading (Ueberschär et al., 2019), when compared to running on soft grass or dirt surfaces. However, more recent studies observed no difference between running surfaces (Milner et al., 2020; Miltko et al., 2022). In addition, a recent systematic review comparing treadmill and overground running concluded that the running kinematics and kinetics are comparable across modalities (Van Hooren et al., 2020). To date, however, the literature which has examined the biomechanical effects of running on different overground surfaces has not been comprehensively summarised. The aim of this study was to therefore systematically review and critically appraise the literature regarding surface compliance on overground running biomechanics.
Methods

A systematic review of the literature was undertaken using the methodology described by the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) Protocols Statement as seen in Figure 1 (Moher et al., 2015). An extensive literature search was carried out using the following electronic databases: CINAHL, EMBASE and EBSCO. The search was built with the help of a research librarian (DM) based on anchoring terms from the following categories: running, overground surface, biomechanics, kinematics, tibial acceleration, pressure, and force. Search terms were expanded using

![PRISMA flowchart of literature search and selection process (Adapted from Moher et al., 2015).]
a vast list of alternative terminologies, truncations and abbreviations. Additional relevant publications were also sought out by retrospective bibliography searches of all included studies and by manually searching for other publications from authors of overground running studies that were identified in the search. Final search was conducted on the 13 November 2022.

**Study selection**

Three independent reviewers (CM, SM, KO) completed an initial title screen to remove any irrelevant papers. The eligibility criteria (Table 1) were designed based on the PICO model (Population, Intervention, Comparison, Outcome). Pilot testing of the exclusion criteria was conducted using a subset of 100 abstracts screened by all three reviewers and the reasons for exclusion were documented. A Cohen’s Kappa of 0.78 was reached and deemed sufficient to conduct the full screening process. Upon completion of the screening process, full texts were reviewed for inclusion and all reasons for article exclusion were recorded. If there was any uncertainty about inclusion, a fourth reviewer (NF) was consulted until a consensus was reached. The independent reviewers were not blinded to the study authors, institutes or journal titles.

In studies which met all inclusion criteria, information was extracted on (1) participant characteristics, (2) the overground surfaces examined, (3) the footwear used, (4) the experimental protocol, (4) the key outcome measures and (5) main findings of the study.

<table>
<thead>
<tr>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Participants</strong></td>
<td>Healthy, recreationally active and/or competitive running community mean age 18–65</td>
</tr>
<tr>
<td></td>
<td>Running experience of at least a year of running ~10 km/week</td>
</tr>
<tr>
<td></td>
<td>Injury free in the previous 3 months</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>Inclusion of two or more overground running surfaces*</td>
</tr>
<tr>
<td></td>
<td>Running: A flat straight runway of at least 10 m included</td>
</tr>
<tr>
<td></td>
<td>Footwear: cushioned or minimalist running shoes included as a trial arm</td>
</tr>
<tr>
<td><strong>Comparison</strong></td>
<td>Velocity: constant across overground surfaces</td>
</tr>
<tr>
<td></td>
<td>Prospective, observational cross-over study designs examining two or more overground running surfaces</td>
</tr>
<tr>
<td><strong>Outcome Measures</strong></td>
<td>Data presented in each surface condition for at least one of the following: tibial acceleration, vertical ground reaction force, loading rate, lower limb kinematics, ground contact time, EMG</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Analysis Publication</strong></td>
<td>Group mean ± SD included</td>
</tr>
<tr>
<td></td>
<td>Comparative analysis across surface conditions</td>
</tr>
<tr>
<td></td>
<td>Date &gt;1990</td>
</tr>
<tr>
<td></td>
<td>Published in peer-reviewed journal</td>
</tr>
<tr>
<td></td>
<td>Full-text available in English</td>
</tr>
</tbody>
</table>

*Surface Conditions: A- Asphalt, AG- Artificial grass/turf, AS- Acrylic surface, C- Concrete, D- Dirt, EVA- Ethylene-Vinyl Acetate, Gr- Gravel, NG- Natural grass, SG- short grass, LG- Long grass, R- Rubber, Ta- Tartan/synthetic track, Tr- Treadmill, P- Pavement, W- Woodchip trail, SS—Sport-Specific.
If available, group mean and standard deviation (SD) data were extracted for key outcome measures including peak tibial acceleration, peak vGRF/plantar pressure, loading rate, ground contact time and discrete joint kinematics.

**Quality appraisal**

To appraise the quality of the included full texts, a Downs and Black (D&B) Checklist was employed (Downs & Black, 1998). This tool has previously been used to evaluate non-randomised controlled trials (non-RCTs) (Taylor et al., 2021). Three independent reviewers (CM, SM, KO) conducted the quality appraisal, and any disagreement was discussed with a fourth reviewer (NF) until consensus was reached. A Cohen’s Kappa of 0.81 was reached. The D&B Checklist is a 27-item list that evaluates methodological strengths and weaknesses of articles based on the categories of (1) Reporting, (2) Internal Validity (Bias), (3) Internal Validity (Confounding), (4) External Validity and (5) Power (Downs & Black, 1998). Statistical power was assessed for each study using the G*Power Application (Faul et al., 2007). 0 marks were awarded for an estimated power < 70%, 1 mark for 70–80%, 2 marks for 80–85%, 3 marks for 85–90%, 4 marks for 90–95% and 5 marks for an estimated power > 95%. Maximum score on the D&B checklist was therefore 32, with overall study quality then expressed as a percentage. The following criteria were used to categorise studies by overall quality (Hamdan et al., 2020): excellent (91%–100%), good (71%–90%), fair (51%–70%) and poor (0%–50%).

**Duplicate data published**

Inclusion of duplicate data risks biasing the results of any meta-analysis (Kwon et al., 2015) and therefore, efforts were made to remove these from subsequent meta-analyses. Upon full-text review of the included studies, it became apparent that duplicate data were likely to have been published across several of the studies. Identical anthropometrics and GCT data are presented in both Fu (2013) and Fu et al. (2015), despite only three participants being included in 2013 and 13 participants included in 2015. The only difference in plantar pressure data presented in Wang et al. (2012) and Hong et al. (2012) appears to be the addition of one more participant (16 ran in Wang’s study while 15 ran in Hong’s). Correspondence with authors confirmed that 10 of the 18 participant data sets presented in Abdul Yamin et al. (2021) are the same data presented in the group’s previous 2017 study (Abdul Yamin et al., 2017). These discoveries are of concern, highlighting a need for greater rigour in the peer-review process and greater accountability to ethical standards for research reporting.

Duplicate data were identified if the following criteria were met: (1) two or more of the authors were common to both articles, (2) the methodology was largely identical and (3) the same outcome variables were reported in both papers. In such instances, authors were contacted to confirm the inclusion of duplicate data. In the event of no response, only data from the highest quality paper (D&B score) were included in any subsequent meta-analyses.
Statistical analysis

An overall effect size was calculated by comparing the difference in peak resultant tibial acceleration, peak vGRF/planter pressure, loading rate and ground contact time on soft and hard overground surfaces. Where data were collected on more than two surfaces, data recorded on the softest and hardest surfaces were extracted for statistical comparison, with the following logic implemented (rubber < grass < woodchip < dirt < tartan < asphalt < concrete). This logic was applied based on measurements of surface compliance using ASTM F2117-10 protocol. In the case of resultant tibial acceleration and ground contact time, group mean and SD on soft and hard surfaces were used to calculate the effect size. In the case of Peak vGFR/Plantar Pressure and Loading Rate, SMD was calculated as kinetic data were in some instances presented in kPa and in others relative to body mass (Tillman et al., 2002).

A random-effects model was used for each meta-analysis. The random-effects model considers these additional sources of between-study variability as well as within-study variability. The I² statistic was used to estimate the percentage of variability across the pooled estimates attributable to heterogeneity beyond chance (where I² of 0% to 25% = low, I² of 26% to 75% = moderate, I² of 76% to 100% = high, severity of between-study heterogeneity).

Where study protocols involved running at more than one running velocity or more than one participant cohort was included in the experiment, the data were averaged to provide a single group mean and SD for each surface condition. However, where study protocols involved running in more than one footwear type, each footwear type was treated separately. This was done on the assumption that cushioning of each footwear type is likely to influence the effect of surface compliance to an extent that it was worth considering this variable separately within each meta-analysis. The potential for bias was evaluated visually by constructing a funnel plot to display the precision of the estimate of the effect size against the estimate of the effect size.

All statistical analysis was conducted using Revman 5.4 (version 5.4, the Nordic Cochrane Centre, Copenhagen, Denmark). All results are presented as mean (SD) unless otherwise stated. A P-value of 0.05 was considered statistically significant.

Results

10,242 references were found from the initial search of online databases. After deduplication, there were 3552 referenced studies to screen. Following the title and abstract screening, 48 studies remained. The bibliographies of these studies were retrospectively screened for missing articles of interest (n = 20), which were also included in full-text review (n = 68). The full-text review yielded 25 articles which met all inclusion criteria (Table 1), 22 of which were later included in quantitative synthesis (Figure 1).

Quality appraisal scores

Results from the modified Downs and Black Quality Checklist are presented in Table 2. The majority of the studies (n = 16) were categorised into the fair range (51–70%). Of the nine remaining studies, five were categorised into the poor range (<50%) and four studies
were categorised into the good range (71–90%). Tessutti et al. (2012) achieved the highest score 25/32 (78%) while Creagh et al. (1998) achieved the lowest score 13/32 (41%). The majority of studies (n = 18) were statistically underpowered (<70%), highlighting the need for quantitative synthesis.

**Participant characteristics**

492 participants were included in this systematic review. A descriptive summary of participant characteristics is presented in Table 3. The majority of participants were male (n = 342; 70%), with 13 studies analysing data solely from male participants. Two studies collected data solely from a female cohort (Creagh et al., 1998; Dixon et al., 2000) and nine studies included both males and females. 21 of the studies reported participant age (n = 401, 28 ± 4 yrs); 22 reported height (n = 442; 173 ± 6 cm) and 23 weight (n = 448; 67 ± 8 kg). Two studies included sub-cohort comparisons defined by running experience or competitive level. Hébert-Losier et al. (2015) compared elite (n = 7) and amateur (n = 7) male orienteers running on three different surfaces. Boey et al. (2017) compared untrained (<2hrs/week; n = 12), recreational (10–30 km; n = 12) and well-trained (>50 km; n = 11) participants.

The natural foot-strike pattern of runners at ground contact in the gait cycle is known to significantly influence loading rates and GRF in the initial stance phase (Lieberman et al., 2010), and is a potential confounding factor in studies examining the external environment on running biomechanics. Of the 25 studies, 12 clearly identified the foot strike pattern of participants with three studies recruiting forefoot
<table>
<thead>
<tr>
<th>Author</th>
<th>Participants (training volume)</th>
<th>Surfaces (measured)</th>
<th>Intervention (familiarisation)</th>
<th>Footwear</th>
<th>Outcome Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdul Yamin et al. (2017)</td>
<td>Male recreational runners (n = 10; 24 ± 1 y; 172 ± 3 cm; 67 ± 7 kg) [NS]</td>
<td>C, AG &amp; R [Y]</td>
<td>Running at SSS on 10 m runway. Three trials on each surface and footwear condition. Only heel strikes included in analysis.</td>
<td>BF, Heeled (NS)</td>
<td>GRF, Spatiotemporal Kinematics (GCT)</td>
<td>GRF &amp; GCT decrease with increasing surface stiffness (only in BF)</td>
</tr>
<tr>
<td></td>
<td>Male university students. Rearfoot runners (n = 18; 24 ± 1 y; 172 ± 3 cm; 67 ± 7 kg). [NS]</td>
<td>C, AG &amp; R [Y]</td>
<td>Running at SSS on 7 m runway. NS trials on each surface and footwear condition. Only heel strikes included in analysis.</td>
<td>Heeled (NS)</td>
<td>Lower Limb Kinematics (Foot)</td>
<td>Significant differences in GCT and MLA angle across surfaces (P = 0.9 and a-priori set at P &lt; 0.1)</td>
</tr>
<tr>
<td>Boey et al. (2017)</td>
<td>Untrained (n = 12; 6 M), recreational (n = 12; 6 M) and well-trained runners (n = 11; 6 M) [Untrained &lt; 2 h.wk(^{-1}); well-trained &gt; 50 km.wk(^{-1})]</td>
<td>C, Ta &amp; W [Y]</td>
<td>Running at SSS and 3.1 m/s for 90 m. First and last 10 m excluded; 2 trials on each surface.</td>
<td>NS</td>
<td>Tibial Acc.</td>
<td>Tibial Acc. significantly lower on W vs C &amp; R</td>
</tr>
<tr>
<td>Creagh et al. (1998)</td>
<td>Female club-level runners (n = 9; 33 ± 9 y)</td>
<td>C, SG &amp; LG [NS]</td>
<td>Running at 4.3 m/s for NS distance; two trials on each surface. Two strides digitised per trial</td>
<td>NS</td>
<td>Spatiotemporal kinematics (GCT, stride length), Lower Limb Kinematics (Knee &amp; Hip), vertical oscillation.</td>
<td>Step length decreased, hip and knee lift increased with decreasing surface compliance (LG &gt; SG &gt; C)</td>
</tr>
<tr>
<td>Dixon et al. (2000)</td>
<td>Female well-trained rearfoot runners (n = 6; 56 ± 3 kg). [NS]</td>
<td>A, R &amp; AS [NS]</td>
<td>Running at 3.3 m/s ± 5% on surfaces 15 m length; 10 trials/surface</td>
<td>Heeled (Adidas Galaxy II)</td>
<td>GRF, Loading Rate, Lower Limb Kinematics (ankle &amp; knee angle at IC and peak angle during stride)</td>
<td>Significant decrease in loading rate of peak impact force on R vs A</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author</th>
<th>Participants (training volume)</th>
<th>Surfaces (measured)</th>
<th>Intervention (familiarisation)</th>
<th>Footwear</th>
<th>Outcome Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dolenec et al. (2015)</td>
<td>Recreational rearfoot runners (n = 8; 6 \text{ M}; 25 \pm 2 \text{ yr}; 177 \pm 7 \text{ cm}, 75 \pm 12 \text{ kg}) ([\text{NS}])</td>
<td>A, Gr &amp; NG</td>
<td>Running at SSS &amp; 4.5 m/s on surfaces 30 m; three trials/surface ([\text{NS}])</td>
<td>NS</td>
<td>EMG, Spatiotemporal Kinematics (velocity, GCT, and step rate)</td>
<td>A requires significantly greater muscle activation of TA and medial GM vs NG. Gr requires significantly greater PB activation vs NG.</td>
</tr>
<tr>
<td>Fu (2013)</td>
<td>Male recreational runners (n = 3, \text{ age 24 } \pm 1; \text{ height 174 } \pm 6 \text{ cm}; 66 \pm 5 \text{ kg}) ([&gt;20 \text{ km.wk}^{-1}])</td>
<td>C, NG, R</td>
<td>Running at 3.3–3.5 m/s on 15 m surfaces; three trials/surface; 10 strides collected/surface</td>
<td>NS</td>
<td>Planter Pressure, Spatiotemporal Kinematics (GCT)</td>
<td>1\textsuperscript{st} peak pressure during stance phase when running on C is significantly greater than NG &amp; R</td>
</tr>
<tr>
<td>Fu et al. (2015)</td>
<td>Male recreational rearfoot runners (n = 13, 24 \pm 1 \text{ yr}; 174 \pm 6 \text{ cm}; 66 \pm 5 \text{ kg}) ([20 \pm 5 \text{ km.wk}^{-1}])</td>
<td>C, NG, Ta, Tr &amp; EVA-Tr ([\text{Y}])</td>
<td>Running at 3.3 m/s on surfaces 30 m with first and last 7.5 m excluded; three trials/surface; 10 strides collected/surface</td>
<td>Minimal (Shanghong Shoes)</td>
<td>Plantar pressure, Loading Rate, Tibial Acc., Spatiotemporal Kinematics (GCT)</td>
<td>EVA-Tr significantly decreased 1\textsuperscript{st} peak pressure &amp; pressure-time integral of impact phase vs C</td>
</tr>
<tr>
<td>Garcia et al. (2021)</td>
<td>Recreational runners (n = 15; 3 \text{ M}; 28 \pm 9 \text{ yr}; 165 \pm 6 \text{ cm}, 66 \pm 13 \text{ kg}) ([\geq \text{3 runs.wk}^{-1}])</td>
<td>D, Gr, P</td>
<td>Running at SSS on surfaces 50 m with first and last 10 m excluded; four trials/surface</td>
<td>Heeled (Saucony Jazz)</td>
<td>Spatiotemporal Kinematics (GCT, stride rate), Tibial Acc., Shock Attenuation</td>
<td>No significant difference found for Tibial Acc. among surfaces</td>
</tr>
<tr>
<td>Greenhalgh et al. (2012)</td>
<td>Male field hockey athletes (n = 9, 21 \pm 2 \text{ yr}, 176 \pm 7 \text{ cm}; 78 \pm 12 \text{ kg}) ([\text{NS}])</td>
<td>C &amp; SS</td>
<td>Running at 3.3 m/s &amp; 5.0 m/s on surfaces, 10 m, NS trials/surface; six footfalls recorded/surface/footwear/speed</td>
<td>Tibial Acc.,</td>
<td>3 \times Hockey; 1 \times Soccer; 1 \times Heeled running shoe (Saucony Jazz)</td>
<td>Significant surface &amp; speed effect, and interaction effect Tibial Acc. significantly higher running at 5.0 m/s on C vs AG No difference between surfaces when running at 3.3 m/s</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author et al. (Year)</th>
<th>Participants (training volume)</th>
<th>Surfaces (measured)</th>
<th>Intervention (familiarisation)</th>
<th>Footwear</th>
<th>Outcome Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollis et al. (2019)</td>
<td>Recreational rearfoot runners ($n = 15; 7 M; 20 ± 3 yr; 170 ± 8 cm; 65 ± 12 kg$) [$27 ± 16 km.wk$^{-1}$]</td>
<td>Ta, NG</td>
<td>Running at SSS of RPE 3–4 and 5–6 on surfaces, 1600 m, After each 400 m lap RPE recorded. Surfaces tested on separate days</td>
<td>Participants own running shoes</td>
<td>Lower Limb Kinematics (foot), Spatiotemporal Kinematics (GCT, step count), Tibial Acc.</td>
<td>GCT was significantly reduced on Ta vs NG during greater RPE runs. Greater max pronation velocity on Ta vs NG. Pronation excursion, braking and impact acceleration were significantly greater on Ta vs NG. Cycle time decreased on Ta vs NG.</td>
</tr>
<tr>
<td>Hébert-Losier et al. (2015)</td>
<td>Male elite &amp; amateur orienteers ($n = 14; 28 ± 7 yr; 183 cm ±6 cm; 74 ± 7 kg$) [Elite ≥10 h.wk$^{-1}$ Amateur 3.6–10 h.wk$^{-1}$]</td>
<td>Road, NG, W</td>
<td>Running at 3.8 m/s on surfaces 20 m; three trials/surface</td>
<td>NS</td>
<td>Spatiotemporal Kinematics (GCT), Lower Limb Kinematics (ankle, knee, hip ROM and peak angle)</td>
<td>Significantly lower velocity, increased step length &amp; cycle time, increased knee extension at foot strike, decreased peak hip flexion &amp; dorsiflexion at stance phase, and increased vertical pelvic motion comparing W vs Road</td>
</tr>
<tr>
<td>Hong et al. (2012)</td>
<td>Male well-trained rearfoot runners ($n = 16; 23 ± 2 yr; 170 ± 5 cm; 64 ± 10 kg$) [&gt;10 km.wk$^{-1}$]</td>
<td>C, NG &amp; Tr</td>
<td>Running at 3.8 m/s on surfaces for 30 m with 5 m measurement zone; five trials/surface</td>
<td>S (TN600-neutral, ASICS)</td>
<td>Plantar Pressure</td>
<td>Tr significantly lower max plantar pressure and force for total foot &amp; 2 toe regions vs C &amp; NG Tr significantly lower average plantar force of medial forefoot &amp; 2 toe regions vs C &amp; NG</td>
</tr>
<tr>
<td>Milner et al. (2020)</td>
<td>Recreational rearfoot runners ($n = 19; 31 ± 6 yr; 170 ± 8 cm; 69 ± 12 kg$) [&gt;10 miles.wk$^{-1}$]</td>
<td>C, NG, Tr &amp; Lab</td>
<td>Running at 3.0 m/ ± 5% on surfaces 1 min five trials/surface</td>
<td>Participants own running shoes</td>
<td>Tibial Acc.</td>
<td>Peak positive acceleration decreased in Lab &amp; Tr vs C &amp; NG</td>
</tr>
<tr>
<td>Miltko et al. (2022)</td>
<td>Recreational rearfoot runners ($n = 12; 30 ± 6 yr; 170 ± 10 cm; 65 ± 10 kg$) [33 ± 20 km.wk$^{-1}$]</td>
<td>C, NG, Ta, Tr</td>
<td>Running at SSS and 20% faster on surfaces for 50 m with 35 m measurement zone; NS the number of trials</td>
<td>S (1080, New Balance)</td>
<td>GRF, Loading Rate, Tibial Acc. (measured using instrumented insoles)</td>
<td>No interaction or surface effects were observed. The faster speed produced greater peak PTA (+19.2%; $p &lt; 0.001$), resultan $	ext{PTA}$ (+20.7%; $p &lt; 0.001$), peak vGRF (+6.6%; $p = 0.002$) and IVLR (+16.5%; $p &lt; 0.001$).</td>
</tr>
</tbody>
</table>

(Continued)
Table 3. (Continued).

<table>
<thead>
<tr>
<th>Author</th>
<th>Participants (training volume)</th>
<th>Surfaces (measured)</th>
<th>Intervention (familiarisation)</th>
<th>Footwear</th>
<th>Outcome Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tessutti et al. (2010)</td>
<td>Recreational runners (n = 44; 32 M; 177 ± 6 cm; 76 ± 11 kg; 12 F; 163 ± 5 cm; 58 ± 4 kg) [36 ± 13 km.wk⁻¹]</td>
<td>NG &amp; A</td>
<td>Running at 12 km/h on surfaces 40 m with first and last 10 m excluded three trials/surface</td>
<td>S (RAINHA)</td>
<td>Plantar Pressure, Lower Limb Kinematics (Foot)</td>
<td>Peak pressure significantly greater on A vs NG; central and lateral rearfoot &amp; lateral forefoot. GCT increased in central forefoot on NG vs A. NG resulted in lower loads in rearfoot and forefoot vs A</td>
</tr>
<tr>
<td>Tessutti et al. (2012)</td>
<td>Recreational runners (n = 47; 34 M; 178 ± 6 cm; 74 ± 11 kg; 13 F; 159 ± 5 cm; 53 ± 4 kg) [38 ± 13 km.wk⁻¹]</td>
<td>A, C, NG &amp; R</td>
<td>Running at 12 km/h on surfaces with first and last 10 m excluded three trials/surface</td>
<td>Heeled (RAINHA)</td>
<td>Plantar Pressure, Lower Limb Kinematics (Foot)</td>
<td>NG decreased peak pressures in rearfoot &amp; forefoot. GCT was greater on R vs C in rearfoot &amp; forefoot A &amp; C similar for all plantar variables and pressure zones</td>
</tr>
<tr>
<td>Tillman et al. (2002)</td>
<td>Male recreational runners (n = 11; 22 ± 3 yr; 177 ± 8 cm; 74 ± 9 kg) [NS]</td>
<td>A, C, NG &amp; R</td>
<td>Running at SSS on surfaces 15 m three trials/surface</td>
<td>S (Shoes provided but NS)</td>
<td>Planar Pressure, Loading Rate, Spatiotemporal Kinematics (GCT)</td>
<td>No significant difference detected among surfaces for reaction forces, GCT or impulse</td>
</tr>
<tr>
<td>Ueberschär et al. (2019) Part 2</td>
<td>Well-trained male runners (n = 8; 29 ± 6 yr; 179 ± 6 cm; 73 ± 6 kg) [5 km PB of &lt;19 min]</td>
<td>Ta, AG &amp; A</td>
<td>Running at 12, 14, 16 &amp; 18 km/h on surfaces 3 mins with first and last 30s excluded two trials/surface</td>
<td>NS</td>
<td>Tibial Acc.</td>
<td>Only AG significantly decrease Tibial Acc. vs Ta &amp; A</td>
</tr>
<tr>
<td>Wang et al. (2012)</td>
<td>Well-trained male rearfoot runners (n = 16; 22 ± 2 yr; 170 ± 3 cm; 63 ± 10 kg), [&gt;20 km.wk⁻¹]</td>
<td>C, NG, Tr</td>
<td>Running at 3.8 m/s on surfaces 30 m with 5 m measurement zone. five trials/surface</td>
<td>S (TN600-neutral, ASICS)</td>
<td>Plantar Pressure</td>
<td>NG significantly decreased max plantar pressure of total foot, lateral midfoot, central forefoot, and lateral forefoot. NG increased GCT at central forefoot and lateral forefoot.</td>
</tr>
<tr>
<td>Willwacher et al. (2014)</td>
<td>Recreational runners (n = 39; 20 M; 24 ± 2 yr; 181 ± 5 cm; 74 ± 6 kg; 19 F; 26 ± 4 yr; 171 ± 6 cm; 60 ± 8 kg) [&gt;10 km.wk⁻¹]</td>
<td>Ta, Ta-Turf, Ta-EVA, Ta-EVA-Turf [Y]</td>
<td>Running at 3.5 m/s ± 5% on surfaces 25 m with midpoint force platform five trials/surface/footwear</td>
<td>BF &amp; S (Brooks Glycerin)</td>
<td>Lower Limb Kinematics at IC (Foot, thigh, pelvis, shank, ankle, knee and hip at initial contact)</td>
<td>Significant shoe x surface interactions for ankle and knee angle at IC; and foot, shank and pelvis orientation at IC. In BF trials on harder surfaces, greater knee extension and plantarflexion at IC was observed. The opposite occurred in shod trials on harder surfaces.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Author</th>
<th>Participants (training volume)</th>
<th>Surfaces (measured)</th>
<th>Intervention (familiarisation)</th>
<th>Footwear</th>
<th>Outcome Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willwacher et al.</td>
<td>Male recreational runners (n=37; 20 M; 24 ± 2 yr; 181 ± 5 cm; 74 ± 6 kg; 17 F; 26 ± 4 yr; 171 ± 6 cm; 60 ± 8 kg) (&gt;10 km.wk⁻¹)</td>
<td>Ta, Ta-Turf, Ta-EVA, Ta-EVA-Turf</td>
<td>Running at 3.5 m/s ± 5% on surfaces 25 m NS trial number [Y]</td>
<td>BF &amp; S (Brooks Glycerin)</td>
<td>Joint moment (hip, knee &amp; ankle), Spatiotemporal kinetics (GCT)</td>
<td>BF + harder surfaces resulted in runners landing more plantarflexed</td>
</tr>
<tr>
<td>Zhang et al.</td>
<td>Male recreational forefoot runners (n=26; 28 ± 7 yr; 173 ± 4 cm; 68 ± 10 kg) [15 ± 12 km.wk⁻¹]</td>
<td>AG, R &amp; C [Y]</td>
<td>Running at 3.3 m/s ± 0.2 m/s on surfaces 15 m three trials/surface [Y]</td>
<td>S (Sortiemagic Rp 4, ASICS)</td>
<td>Planar Pressure</td>
<td>Surface effects stronger in BF vs S</td>
</tr>
<tr>
<td>Zhou et al.</td>
<td>Male recreational forefoot runners (n=31; 29 ± 7 yr; 174 ± 5 cm; 68 ± 10 kg) [≥10 km.wk⁻¹]</td>
<td>AG, C &amp; R [Y]</td>
<td>Running at 3.3 m/s ± 0.2 m/s on surfaces 15 m with two force platforms at 8 m x five trials</td>
<td>S (Sortiemagic Rp 4, ASICS)</td>
<td>Lower Limb Kinematics, Joint moment, Loading Rate (Hip, Knee &amp; Ankle)</td>
<td>Decreased pressure-time integral in central forefoot and lateral forefoot Kinematics; Max knee flexion angle greater when running on AG vs R &amp; C Kinetics; peak vGRF higher on R vs C, lower limb joint moments decreased on R vs C</td>
</tr>
<tr>
<td>Zeng et al.</td>
<td>Male recreational forefoot runners (n=31; 29 ± 7 yr; 174 ± 5 cm; 68 ± 10 kg) [34 ± 25 km.wk⁻¹]</td>
<td>AG, R, C [Y]</td>
<td>Running at 3.3 m/s ± 5% on surfaces 15 m 3 trials/surface [Y]</td>
<td>S (Sortiemagic Rp 4, ASICS)</td>
<td>GRF, Lower limb Stiffness</td>
<td>No difference in vertical loading &amp; joint stiffness; AG increased knee joint angle displacement vs R &amp; C; Peak knee moment decreased on C vs AG; Peak ankle moment decreased on R vs AG &amp; C; Max GRF decreased in C.</td>
</tr>
</tbody>
</table>

**Abbreviations:*** Details of the demographics of included participants; F- Female, M- Male. Surface characteristics; A- Asphalt, AG- Artificial grass/turf, AS- Acrylic surface, C- Concrete, D- Dirt, EVA- Ethylene-Vinyl Acetate, Gr- Gravel, NG- Natural grass, SG- Short grass, LG- Long grass, R- Rubber, Ta- Tartan/synthetic track, Tr- Treadmill, P- Pavement, W- Woodchip trail, SS—Sport-Specific. Study intervention; SSS—self-selected speed, RPE—Rate of Perceived Exertion. Footwear; BF- Barefoot, S- Standard. Outcome Measures; GCT- Ground Contact Time, GRF- Ground Reaction Force, MLA- Medial Longitudinal Arch, PFS- Plantar Fascia Strain, IC- Initial Contact, Main Findings; TA—Tibialis Anterior, GM- Gastrocnemius. Others; NS- Not Specified, Y- Yes.
strikers (Zeng et al., 2021; Zhang et al., 2020; Zhou et al., 2021) and nine recruiting rearfoot strikers. The remaining 13 studies did not specify the foot strike pattern of participants (see Table 3).

**Study characteristics**

All studies included at least two of the following overground surfaces; concrete (i.e., sidewalk), asphalt (i.e., road), artificial grass/turf, natural grass (long and short), rubber/artificial track, woodchip trail, acrylic surface, gravel, dirt and tartan track material (see Table 3). Concrete was the most frequently studied surface (17 studies). A tartan synthetic track material was used in seven studies, while a rubber synthetic track was used in nine studies. Natural grass was used in 52% of the studies (n = 13) while artificial grass/turf was used in six studies. In total, 18 of the included studies provided a familiarisation period on each surface prior to data collection (see Table 3). Target velocity (m/s or km/h) was specified in 19 studies, four of which included multiple velocities (Boey et al., 2017; Dolenec et al., 2015; Greenhalgh et al., 2012; Ueberschär et al., 2019). The remaining six studies defined running velocity to be ‘self-selected’ per participant.

Totally, 20 studies specified the use of only one footwear type while the remaining five studies used multiple footwear conditions across overground surfaces (see Table 3). Abdul Yamin et al. (2017) compared barefoot, minimalist running shoes and heeled running shoes across different surfaces. Greenhalgh et al. (2012) compared five separate shoes; a standard running shoe (Saucony Jazz), a soccer-specific AstroTurf shoe (Umbro Soccer) and three different hockey-specific shoes (Asics Gel Lethal, Gryphon Venom, Gryphon Viper). Willwacher et al. (2014, 2022) compared barefoot vs standard running shoes, while Abdul Yamin et al. (2021) compared heeled vs. minimal running shoes. Totally, 11 studies did not specify the make/brand of shoes used.

Each of the included studies measured at least one of the following outcome variables across surface conditions; peak tibial acceleration (n = 7), GRF (n = 5), peak plantar pressure (n = 8), loading rate (n = 5), lower limb kinematics (n = 10), spatiotemporal kinematics (n = 11) and lower limb muscle activity (n = 1). In addition, three studies combined their GRF data and joint kinematic data via inverse dynamics to quantify joint moment during the gait cycle (Willwacher et al., 2022; Zeng et al., 2021; Zhou et al., 2021).

**Quantitative comparison**

Forest plots of the quantitative synthesis of study data comparing outcome variables between soft and hard surfaces are presented in Figure 2. Random effects analysis found that peak tibial acceleration was significantly lower when running on softer surfaces (P = 0.01, Z = 2.51; SMD=–0.8; 95% CI = −1.42 to −0.18), see Figure 2(a). However, neither vGRF/plantar pressure (P = 0.18; Z = 1.34; SMD = −0.14; 95% CI = −0.35 to 0.07) nor loading rate (P = 0.28; Z = 1.08; SMD = −0.42; 95% CI = −1.17 to 0.34) were significantly different when comparing hard and soft surfaces. In addition, surface compliance had no effect on GCT (P = 0.86; Z = 0.18; MD = 0.76; 95% CI = −7.53 to 9.06). Heterogeneity between studies was low for peak tibial acceleration (n = 7; I2 = 0%) and vGRF/Plantar pressure (n = 9; I2 = 0%), moderate for GCT (n = 9; I2 = 75%) and high for loading rate (n = 5; I2 = 76%).
The purpose of this review was to systematically appraise the literature investigating surface compliance and its effect on overground running biomechanics. We identified 25 studies which compared two or more overground running surfaces, the majority (n = 16) of which were of fair quality. There were also several studies identified where...
duplicate data appeared to be published. The main finding of our meta-analyses was that running on softer surfaces significantly decreases tibial acceleration; however, no effect of surface compliance was observed for any other kinetic or kinematic variables evaluated. Due to the inconsistent reporting of lower limb kinematics or joint moment, a meta-analysis of these variables was not possible. However, they do provide helpful insight into how surface compliance might affect running patterns. Three studies highlighted the interaction between footwear, surface and biomechanics (Abdul Yamin et al., 2021; Greenhalgh et al., 2012; Willwacher et al., 2022), with footwear types interacting with surface compliance in contrasting ways (see Figure 2). Lastly, this systematic review revealed a distinct lack of consistency in terms of methodological reporting. Future studies should ensure an adequate description of key study features including a quantitative measurement of surface compliance, characteristics of the running cohort, footwear used, familiarisation with surfaces, and the velocity and duration of data collection trials. In addition, consistent reporting of key biomechanical outcome variables would provide greater insight for future reviews.

Effect of surface compliance on tibial acceleration

Stress fractures account for approximately 20% of all RRI (Arnold & Moody, 2018). The tibia is the bone most likely to be fractured (35–56% of all stress fracture injuries) with a higher prevalence observed in female runners (Matheson et al., 1987; Romani et al., 2002). MTSS, a precursor to stress fracture, accounts for up to 16% of all RRIs and has been suggested as a contributing factor in up to 50% of all lower limb injuries (Craig, 2008; Kortebein et al., 2000; Yates & White, 2004). A relationship exists between peak tibial acceleration measured during running and risk of tibial stress fracture. In an observational case-controlled study, peak tibial acceleration successfully predicted with 70% of those with a history of TSF, via binary logistic regression (Milner et al., 2006). Our meta-analysis of 11 studies and 130 participants demonstrated a significant reduction in tibial acceleration ($Z = 2.58, P = 0.01$) when running on softer surfaces, therefore supporting the common recommendation of overstretch avoidance through running on softer surfaces (Warden et al., 2014). Conversely, running on hard surfaces increases mechanical strain on this bone (Sheerin et al., 2019; Tenforde et al., 2020), potentially increasing long-term risk of tibial stress injury.

Despite low heterogeneity within the meta-analysis for peak tibial acceleration, it is worth noting that considerable variance in the location of accelerometer placement on the tibia does exist within the literature and may greatly influence the results (Lucas-Cuevas et al., 2017; Sheerin et al., 2019). Distal placement has previously been shown to result in greater tibial acceleration compared to a more proximal placement (Lucas-Cuevas et al., 2017). Such findings highlight that the location of the accelerometer does influence the signal parameters, and thus, a standardised protocol is needed to allow more accurate comparison of data across studies. Within the studies that analysed tibial acceleration and meet inclusion into meta-analysis, three studies (Fu et al., 2015; Greenhalgh et al., 2012; Ueberschär et al., 2019) analysed peak axial tibial acceleration compared to three studies (Garcia et al., 2021; Milner et al., 2020; Miltko et al., 2022) that analysed both peak axial and resultant tibial acceleration. It is unlikely that this difference
in investigation impacted our findings. Garcia et al. (2021), Milner et al. (2020) and Miltko et al. (2022) found no statistical difference between peak axial and resultant tibial acceleration. While resultant acceleration was consistently higher than peak axial acceleration, both displayed the same pattern across overground and laboratory conditions (Milner et al., 2020).

**No effect on loading rate**

Tibial acceleration has been strongly correlated with vertical loading rates measured on an instrumented force platform (Hennig et al., 1993; Lafortune et al., 1995; Tenforde et al., 2020). Furthermore, in the observational case-control study conducted by Milner et al. (2006), only tibial acceleration and loading rates were associated with risk of TSF. It is reasonable therefore to expect that a quantitative synthesis of loading rate data would mirror the results observed for tibial acceleration. However, our meta-analysis of 5 studies and 68 participants found no significant difference in loading rate ($P = 0.28$; $Z = 1.08$; SMD = $-0.42$; 95% CI = $-1.17$ to $0.34$) when running on hard or soft surfaces. There are two possible reasons for this discrepancy in results for tibial acceleration and loading rates. Firstly, perhaps tibial acceleration is not as strong as a proxy of loading rate as has previously been suggested, and the significant effect observed in tibial acceleration is a false positive. Or perhaps the severe heterogeneity ($I^2 = 76\%$) and low number of included studies ($n = 5$) means, we may be observing a potential false negative in loading rate in the context of the findings from Milner et al. (2006). Substantial methodological heterogeneity within these studies was also apparent. For example, for Fu et al. (2015) and Tillman et al. (2002), loading rates were calculated from the magnitude of the first impact peak, divided by the time at which this occurred. However, Zhou et al. (2021) used the average and steepest slope of the vGRF data of the initial 13% of the stance phase when calculating their loading rates, because they recruited only forefoot strikers who do not produce a discernable impact peak (Blackmore et al., 2016; Zhou et al., 2021). In contrast, Dixon et al. (2000) and Miltko et al. (2022) recruited habitual rearfoot strikers. Miltko et al. (2022) calculated loading rate from the vGRF as the peak magnitude of the first-time derivative in the stance phase. Dixon et al. (2000) used the magnitude and time occurrence of the peak impact force of vGRF data. This heterogeneity likely impacts the recording of loading rate. Schmida et al. (2022) in a prospective observational study found that the calculation method and speed did result in significantly different loading rate values. The maximum slope from 20% to 80% of the vGRF at 4.47 m$\cdot$s$^{-1}$ produced the highest LR estimate and the average slope from initial contact to IP at 2.68 m$\cdot$s$^{-1}$ produced the lowest. However, no association was made between calculation method and subsequent injury risk.

This additional heterogeneity in running technique (forefoot vs rearfoot runners) and its effect on the measure of loading rate likely plays a confounding role in any quantitative synthesis. After all, GRF and by proxy loading rate are both measures of forces acting on the centre of mass of the entire body, while tibial acceleration is a more direct estimate of loading at the tibia (Milner et al., 2006). It may, therefore, be reasonable to conclude that when considering all foot-strike patterns, tibial acceleration serves as better proxy measure of impact mechanics at the tibia, the most common site of bone stress in running (Sheerin et al., 2019). Certainly, the results of our meta-analysis may further
support Milner et al. (2006) who suggested that tibial acceleration may be a more sensitive discriminator of TSF risk than loading rate.

**No effect on peak ground reaction force**

A meta-analysis of nine studies and 177 participants found no significant difference in peak vGRF when running on soft and hard surfaces ($Z = 174; P = 0.18$). This finding supports the recent study which reported that surface characteristics did not affect peak vGRF (Miltko et al., 2022). It is worth noting however, that peak vGRF was measured in some cases using instrumented in-sole (plantar pressure) and in others via 3D force plates. Miltko et al. (2022) made a case for the use of insoles to measure peak vGRF based on previous research investigating its validity and reliability in by Renner et al. (2019). Regardless of the measurement approach, vGRF is generally observed mid stance when the lower limb is fully supporting the body mass overground. Softer surfaces might absorb initial impact forces early in the stance phase (observed in tibial acceleration or loading rate) but ultimately the peak force is likely to end up the same (as is the case comparing barefoot and cushioned running shoes). It is, therefore, unlikely that a change in surface compliance would affect this measure.

**No effect on ground contact time**

A meta-analysis of 9 studies and 211 participants reveals no significant difference in GCT when running on hard or soft surfaces. However, spatiotemporal kinematics like GCT and step rate interact with running kinetics. For example, Gerrard and Bonanno (2018) found that increasing step rate (reducing GCT and stride length) significantly reduced peak plantar pressure. In addition, increasing stride frequency (and by proxy reducing GCT) has been associated with reduced metabolic cost, braking impulses, vertical oscillations of the COM, vGRF and tibial acceleration, and loading at the hip and knee (Anderson et al., 2022; Chumanov et al., 2012; Derrick et al., 1998; Farley & González, 1996; Heiderscheit et al., 2011; Lenhart et al., 2014; Lieberman et al., 2015; Mercer et al., 2003; Morin et al., 2007; Schubert et al., 2014; Seay et al., 2008; Stergiou et al., 2003). Therefore, our meta-analysis reveals no significant difference in GCT when running on hard or soft surfaces. However, GCT and other kinematic variables such as stride frequency and length due have implications when it comes to incidence of RRIs.

**Footwear and strike pattern**

Footwear itself was not treated as a factor, as it isn’t possible to do so in the context of a meta-analysis to include multiple factors. However, for studies that used more than one footwear type, we included each footwear type as a separate entry in the meta-analysis. For example, peak tibial acceleration from each of the three footwear types used in Greenhalgh et al. (2012) were included in that meta-analysis (see Figure 2(a)).

One could argue that the true effect of surface compliance on overground running biomechanics is probably best answered using only studies where participants are running barefoot. However, this approach is less ecologically valid, with most people in the real world running in cushioned shoes. In addition, we
already know that the biomechanics of barefoot running differs substantially from running in cushioned shoes. Habitually barefoot and minimally shod runners adopt a more forefoot strike pattern and higher stride frequency, with corresponding reductions in impact collisions and loading rates (Lieberman et al., 2010). Furthermore, even habitually shod runners almost immediately alter their muscle recruitment patterns when acutely exposed to barefoot running (Fleming et al., 2015). In order to maintain ecological validity, this review therefore only included studies where participants wore cushioned running shoes in at least one trial arm. However, three studies which used multiple footwear types across surface conditions were included in this review. These do allow some closer examination and consideration of the interactions between surface compliance and footwear.

Greenhalgh et al. (2012) investigated the use of three different types of shoes on a concrete and a hockey-specific artificial turf surface. The difference in peak tibial acceleration between those two surfaces was least in the cushioned running shoes (Saucony Jazz) and greatest for the ‘minimally cushioned’ soccer shoes (see Table 3 and Figure 2(a)). This observation suggests that the runner’s sensitivity to changes in surface compliance are in part affected by the level of cushioning in the shoe. Similarly, Abdul Yamin et al. (2021) found that a contrasting relationship existed between surface hardness and GCT when running in different footwear. GCT on soft surfaces was longer when participants ran in minimalist shoes but significantly shorter when running in cushioned heeled shoes (see Figure 2(c)). Willwacher et al. (2022) observed similar findings, where barefoot running produced longer GCTs on a soft surface, compared to shorter GCT when running in Brooks Glycerin running shoes (see Table 3 and Figure 2(c)). This subset of studies highlights the complex interactions between surface compliance and footwear when it comes to changes in running biomechanics.

**Measuring surface compliance**

There was a lack of consistency in the classification of ‘soft’ or ‘hard’ surfaces within the literature. Only 10 of the included studies measured and reported surface compliance with six referencing the use of the ASTM F2117-01 protocol. This protocol involves dropping a standard basketball (size #7 with air pressure of 0.06 k Pa) from a height of 2 m, with five trials required to calculate the surface compliance (in cm rebound height). Tessutti et al. (2012), Willwacher et al. (2014, 2022) Willwacher et al. (2014), Willwacher et al. (2022) and Miltko et al. (2022) did not describe a standardised ASTM protocol. Tessutti et al. (2012) report dropping a rubber ball from 1.5 m and using a high-speed video camera to calculate kinetic energy restored (in ml). Miltko et al. (2022) report dropping a lacrosse ball from a height of 1 m and using a high-speed video to record rebound height. Rebound height was used to measure the collision’s coefficient of restitution [the square root of the quotient of bounce & drop height], a surrogate for surface hardness. In contrast, Willwacher et al. (2014, 2022) measured surface stiffness using a material testing machine which applied 1800N of force over a fixed time and reported stiffness in kN.mm⁻¹. Authors are encouraged in future to use the ASTM F2117-01 protocol in order to maintain consistency in the literature and allow for direct comparison of surface compliance across studies.
Limitations

There are a few limitations to this review that should be considered when interpreting the findings. Notably, the decision to combine peak vGRF and plantar pressure in our meta-analysis (see Figure 2(b)). These two biomechanical measures are inherently different, plantar pressure being force (vGRF) divided by unit area. However, there were only four papers directly measuring vGRF on an instrumented force plate, limiting the strength of any meta-analysis on this important biomechanical measure. Two separate low-powered meta-analyses would have shown no effect, compared to one high-powered meta-analysis showing no effect. Therefore, a SMD approach was used to account for any variance in reported outcome measures that were effectively measuring the same parameter. While this action may be justified, it does highlight an important limitation and consideration when interpreting the findings of our meta-analysis and interpreting results in this area of research.

Secondly, our review observed concerning author data collection and publication practices. Attention is drawn to Fu (2013) and Fu et al. (2015). On the basis of participant characteristics, methodologies and authors, there was likely duplication of data. In which case, only the higher quality paper (Fu et al., 2015) was included for meta-analysis. However, the units of measure for plantar pressure in both studies differ as does the corresponding data. Fu (2013) reports plantar pressure data in N.cm-1.kg-1 (i.e., pressure normalised to body mass). However, Fu et al. (2015) present data in kPa, orders of magnitude lower than other plantar pressure studies. This does limit the strength of our meta-analysis as well as the integrity of author research in this field of study. It is unknown the full impact these studies have had on our comprehensive review, but it is a considerable finding. One which should cause alarm and push the field forward to reform and hypervigilance of author and research publication.

Implications of findings on future research

Following our meta-analysis and systematic review, important study design observations were obtained. Four recommendations can be highlighted and implemented moving forward: (1) improved methodological explanation, (2) standardisation to define surface types (ASTM F2117-01protocol), hardness and/or compliance, (3) increased research vigour and surveillance of republished data, and (4) developing more detailed and focused studies to draw clear conclusions.

The potential impact of such studies would give clarity to researchers on methodological design, such that systematic reviews offer a more comprehensive understanding of the literature. This leads to better knowledge translation to clinicians and the running community.

Implications of findings on clinical practice

Based on our research findings, the following clinical recommendations can be made to runners, coaches and healthcare professionals in sports medicine to help better manage their athletes. (1) Recognition that running and RRI is multifactorial. An individualised approach to your athlete is required to understand the mechanism of RRI and subsequent
management. (2) Importance of workout prescription to include the type of surface. Running on softer surface does appear to lower impact. This should provide coaches and clinicians with confidence that running on softer surfaces may be an effective measure in reducing a risk factor for overstress RRI.

**Conclusion**

The main finding of this systematic review was that running on harder surfaces significantly increases peak tibial acceleration. However, a similar effect was not observed in the corresponding meta-analysis of loading rate data. This contrasting finding is likely due to the severe heterogeneity in studies which measured loading rate, the small number of included studies (n = 5) and differing foot-strike patterns between cohorts. Future studies measuring loading rate in larger cohorts of runners with a rear/midfoot strike pattern would allow for greater insight on the effect of surface compliance on this important outcome variable. Overall, since peak tibial acceleration has been associated with increased risk of tibial stress injuries (Milner et al., 2006), the results of this meta-analysis suggest that the common recommendation of reducing impact stress on the bone by running on softer surfaces is likely justified.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

The author(s) reported there is no funding associated with the work featured in this article.

**ORCID**

Cameron Mitchell [http://orcid.org/0000-0002-3783-0994](http://orcid.org/0000-0002-3783-0994)

**References**


Scheerder, J., Breedveld, K., & Borgers, J. (2015). Who is doing a run with the running boom? In J. Scheerder, K. Breedveld, & J. Borgers (Eds.), *Running across Europe: The rise and size of one of the largest sport markets* (pp. 1–27). Palgrave Macmillan. https://doi.org/10.1057/9781137446374_1


