The effectiveness of exercise interventions to prevent sports injuries: a systematic review and meta-analysis of randomised controlled trials

Jeppe Bo Lauersen,1 Ditte Marie Bertelsen,2 Lars Bo Andersen3,4

ABSTRACT
Background Physical activity is important in both prevention and treatment of many common diseases, but sports injuries can pose serious problems.
Objective To determine whether physical activity exercises can reduce sports injuries and perform stratified analyses of strength training, stretching, proprioception and combinations of these, and provide separate acute and overuse injury estimates.
Material and methods PubMed, EMBASE, Web of Science and SPORTDiscus were searched and yielded 3462 results. Two independent authors selected relevant randomised, controlled trials and quality assessments were conducted by all authors of this paper using the Cochrane collaboration domain-based quality assessment tool. Twelve studies that neglected to account for clustering effects were adjusted. Quantitative analyses were performed in STATA V.12 and sensitivity analysed by intention-to-treat. Heterogeneity (I²) and publication bias (Harbord’s small-study effects) were formally tested.
Results 25 trials, including 26 610 participants with 3464 injuries, were analysed. The overall effect estimate on injury prevention was heterogeneous. Stratified exposure analyses proved no beneficial effect for stretching (RR 0.963 (0.846–1.095)), whereas studies with multiple exposures (RR 0.655 (0.520–0.826), proprioception training (RR 0.550 (0.347–0.869)), and strength training (RR 0.315 (0.207–0.480)) showed a tendency towards increasing effect. Both acute injuries (RR 0.647 (0.502–0.836)) and overuse injuries (RR 0.527 (0.373–0.746)) could be reduced by physical activity programmes. Intention-to-treat sensitivity analyses consistently revealed even more robust effect estimates.
Conclusions Despite a few outstanding studies, consistently favourable estimates were obtained for all injury prevention measures except for stretching. Strength training reduced sports injuries to less than 1/3 and overuse injuries could be almost halved.

INTRODUCTION
Increasing evidence exists, for all age groups, that physical activity is important in both prevention and treatment of some of the most sizable conditions of our time,1–3 including cardiovascular disease, diabetes, cancer, hypertension, obesity, osteoporosis, and depression. Although overall population levels of physical activity is a general concern, increasing levels of leisure time physical activity and sports participation have been reported in some population groups.4 Injuries are virtually the sole drawback of exercise, but may be a common consequence of physical activity and have been shown to pose substantial problems.5–7 Management of sports injuries is difficult, time-consuming and expensive, both for the society and for the individual.8–10 However, sports injury prevention by different kinds of strength training, proprioception exercises, stretching activities, and combinations of these, is accessible to essentially everyone and requires limited medical staff assistance. This adds several interesting aspects regarding the potential dispersion, applicability, and compliance to these programmes.

Most studies on musculoskeletal injuries have focused on one particular intervention, injury type/ location, sport or studied other relatively narrowly defined research questions. This applies to most reviews and meta-analyses as well.11–18 However, Parkkari et al19 described 16 controlled trials in a narrative review. Central concepts of sports injury prevention such as extrinsic (including exposures, environment, equipment) and intrinsic (including physical characteristics, fitness, ability, age, gender, psychology) risk factors and the ‘sequence of prevention’ model of van Mechelen20 were summarised. Aaltonen et al21 presented an overview of all sports injury prevention measures, but as in the literature up until their search in January 2006, the focus of this review was primarily on extrinsic risk factors.22 Recently, and with less restrictive exclusion criteria, Schiff et al23 covered the same topic with additional studies. Aaltonen et al and Schiff et al were unable to obtain full quantification of intervention effect estimates. Steffen et al24 presented a narrative review of acute sports injury prevention written by field experts for each location of injury, but an examination and quantification of specific training exposures and a differentiation of acute and overuse outcome effect estimates is still lacking.

This review and meta-analysis will broaden the scope of previous reviews and meta-analyses on sports injury prevention and focus on the preventive effect of several different forms of physical activity programmes and complement the existing summative literature on extrinsic risk factor reduction. Valuable summary literature exists for both neuromuscular proprioception14 15 and stretching exercises.17 18 However, aggregation of effect estimates and comparison with the effect of strength training and an intervention group with multiple exposures (combining ex strength, proprioception, stretch etc) could reveal new and interesting information, enabling proposals for future directions in the field of sports injury prevention. This study consequently aimed at performing stratified analyses of different injury prevention exercise programmes and additionally provides separate effect estimates for acute and overuse injuries.
MATERIAL AND METHODS

Search strategy and study selection

A review protocol was composed, comprising a priori specification of analyses, inclusion/exclusion criteria, injury definition and search strategy. Injury was defined according to the F-MARC consensus statement for football, merely broadened to fit all forms of physical activity.25 See online supplements eMethods1–3 and eFigure 1 for full injury definition, detailed search entries, study selection description and flow chart. PubMed, EMBASE, Web of Science and SPORTDiscus databases were searched to October 2012 with no publication date restrictions. The search was performed by four blocks of keywords related to prevention, injury and diagnoses, sports, and randomised controlled trials. The searches were customised to accommodate the layout and search methods of each search engine and the application of additional free text words were based on the coverage of subject terms. Reference lists of retrieved articles were hand searched for trials of potential interest and the search was later updated to January 2013.

Search results yielded 3462 hits, which were screened by title to yield 90 titles. After exclusion by abstract, 40 were read in full text and 22 were included. Another three studies were included from reference lists and updated search. Study selection followed a priori-specified inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ Primary prevention</td>
<td>▶ Influencing pathology</td>
</tr>
<tr>
<td>▶ Free of injury at inclusion</td>
<td>▶ Surrogate measures of injury</td>
</tr>
<tr>
<td>▶ Sports/physical activity injuries</td>
<td>▶ Any use of devices (kinesiotaping,</td>
</tr>
<tr>
<td>▶ Randomised controlled trials</td>
<td>insoles, etc)</td>
</tr>
<tr>
<td>▶ Appropriate intervention/control arms</td>
<td>▶ Any means of transportation (bicycles,</td>
</tr>
<tr>
<td>▶ Conducted in humans</td>
<td>motor driven, skies, equestrian, etc)</td>
</tr>
<tr>
<td>▶ Reported in English</td>
<td>▶ Inadequate follow-up</td>
</tr>
<tr>
<td>▶ Peer-reviewed publications</td>
<td></td>
</tr>
</tbody>
</table>

Two reviewers (JBL and DMB) independently assessed the eligibility criteria with subsequent consensus by discussion. If unanimous consensus could not be reached, this was arbitrated by a third person (LBA).

In total, 25 studies were included.26–50

Data extraction

All included studies were assessed using the domain-based evaluation tool recommended by the Cochrane collaboration.51 Two reviewers (JBL and DMB) independently collected the support for judgement and final judgements required consensus from all authors of this paper. If reporting was inadequate or unclear, efforts were made to contact the corresponding authors and ask by ‘open questions’ in order to reduce the risk of overly positive answers. Weighting of studies by quality assessment was considered but not performed, as such appraisals would inevitably involve subjective decisions and no evidence in support of this approach exists.51

Data extraction for total estimate and exposure subgroup estimates covered the primary outcome, defined by each study. Injuries were classified as acute or overuse according to definitions used by each study and proprioception was defined as exercises aiming at improving joint proprioception and/or joint stability. For the outcome subgroups, acute and overuse injuries, we additionally extracted appropriate secondary data from studies where information was available in order to optimise the power of these analyses. Overlapping entities were omitted so no injury was analysed more than once.

The stratification of studies into less heterogeneous exposure subgroups was, with the exception of Beijsterveldt et al.,27 performed after completion of the literature search. Beijsterveldt et al was added from the updated literature search and was unambiguously fitted into the multiple exposures group.

As compliance plays a central role in the robustness of results, sensitivity analyses without studies that neglected to analyse by intention-to-treat were conducted.

During the iterative process of hypothesis generation and preliminary searches the prespecified eligibility criteria were elaborated but not changed. All a priori-specified analyses were performed as planned.

Statistics

Whenever possible, only first-time injuries were taken into account as repeated outcomes are likely to be dependent of each other and therefore would introduce bias. Most studies have analysed by calculation of either RR, injury rate RR or Cox regression RR. When no appropriate effect estimates were reported the study neglected to adjust for clustering effects, we adjusted for clustering effects and calculated a RR. Twelve included studies were not originally adjusted for cluster randomisation. As individuals in clusters potentially lack independence of each other, a regulation of sample size calculations is often required. The equation for cluster adjustment is

\[ IF = 1 + (n - 1) \rho \]

where \( \rho \) is the intracluster correlation coefficient, \( n \) the average cluster size and IF the inflation factor. Effective sample size is calculated by dividing sample size with IF.52 The intracluster correlation coefficient was calculated by

\[ \rho = \frac{s^2_w}{(s^2_w + s^2_c)} \]

where \( s^2_w \) is the within cluster variance of observations taken from individuals in the same cluster and \( s^2_c \) the variance of true cluster means.53 In the nine studies where the corresponding authors did not provide us with sufficient data for \( \rho \) calculation, we achieved this by calculating an average intracluster correlation coefficient based on \( \rho \) values from studies, which were appropriately adjusted for clustering effects.

In order to address reporting bias formally, we sought to test all analyses by the Harbord small-study effect test with a modified Galbraith plot.54 This follows the recommendations by the Cochrane handbook for systematic reviews of interventions and is available in STATAV12.51 55 Effective sample sizes for intervention and control group populations were used for the required binary data input to achieve a cluster-adjusted result for this test.

The heterogeneity for all analyses was assessed by \( I^2 \) and the \( \chi^2 (Q) \) p value. \( I^2 \) is calculated from the Stata given Q value and number of studies (n) by

\[ I^2 = \frac{Q - (n - 1)}{Q} \]

A rough interpretation guide of \( I^2 \) has been proposed by Higgins et al.51

All analyses were computed in STATA V12 by user-written commands described by Egger et al.56 The random effects model was used for the weighting of studies. Statistically heterogeneous
estimates were graphically explored by the metainf command, displaying the influence of each individual study on the effect estimate. These analyses did not reveal conclusive information of particular studies primarily causing the heterogeneity and will not be reported throughout this article.

RESULTS
Study characteristics
Table 1 summarises the characteristics of 25 included studies. A full study characteristics table is available in the online supplement eTable 2. In total 26 610 individuals were included in the analysis and effect estimates were based on 3464 injuries. Thirteen studies were performed on adult participants, 11 studies on adolescents and one study included both.

We contacted nine authors and four supplied clarifying answers with subsequent change in their data or quality assessment. For detailed quality assessments and quality assessment summary see online supplementary eMethods 4, eTable 1 and eFigure 2.

Total estimate
The total effect estimate was RR 0.632 (95% CI 0.533 to 0.750, I²=70% with a χ² p<0.001). Brushoj et al. 28 Eils et al., 30 Gilchrist et al., 34 Holmich et al.36 and Soderman et al. 46 did not report intention-to-treat data. When performing a sensitivity analysis on the 20 studies with intention-to-treat data, an estimate of RR 0.608 (0.503–0.736, I²=74%, χ² p<0.001) was found. A post hoc analysis stratified for age showed RR 0.577 (0.453–0.736, I²=68%, χ² p<0.001) for adolescents and RR 0.683 (0.526–0.885, I²=72%, p<0.001) for adults (figure 1).

<table>
<thead>
<tr>
<th>Table 1 Study characteristics summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>Askling et al.</td>
</tr>
<tr>
<td>Beijsterseldt et al.</td>
</tr>
<tr>
<td>Brushoj et al.</td>
</tr>
<tr>
<td>Coppack et al.</td>
</tr>
<tr>
<td>Eils et al.</td>
</tr>
<tr>
<td>Emery et al.</td>
</tr>
<tr>
<td>Emery and Meeuwisse</td>
</tr>
<tr>
<td>Emery and Meeuwisse</td>
</tr>
<tr>
<td>Gilchrist et al.</td>
</tr>
<tr>
<td>Heidt et al.</td>
</tr>
<tr>
<td>Holmich et al.</td>
</tr>
<tr>
<td>Jantvedt et al.</td>
</tr>
<tr>
<td>LaBella et al.</td>
</tr>
<tr>
<td>Longo et al.</td>
</tr>
<tr>
<td>McGuine and Keene</td>
</tr>
<tr>
<td>Olsen et al.</td>
</tr>
<tr>
<td>Pasanen et al.</td>
</tr>
<tr>
<td>Petersen et al.</td>
</tr>
<tr>
<td>Pope et al.</td>
</tr>
<tr>
<td>Pope et al. 2000</td>
</tr>
<tr>
<td>Soderman et al.</td>
</tr>
<tr>
<td>Soligard et al.</td>
</tr>
<tr>
<td>Steffen et al.</td>
</tr>
<tr>
<td>Walden et al.</td>
</tr>
<tr>
<td>Wedderkopp et al.</td>
</tr>
</tbody>
</table>
Stratified exposure analyses

The strength training estimate including four studies was RR 0.315 (0.207–0.480, I²=0%, χ² p=0.808). All studies in the strength training group were analysed by intention-to-treat. For stratified exposure Forest plots see online supplementary eFigure 4–7.

The pooled effect estimate for six studies with proprioception training as the primary exposure showed a RR of 0.550 (0.347–0.869, I²=66%, χ² p=0.012). Sensitivity analysis of intention-to-treat ruled out Eils et al and Soderman et al, and revealed RR 0.480 (0.268–0.862, I²=71%, χ² p=0.017).

Unlike the above two exposures, the overall estimate for stretching did not prove significant with RR 0.963 (0.846–1.093, I²=0%, χ² p=0.975) based on three studies. All studies in the stretching group were analysed by intention-to-treat.

Stratified outcome analyses

On the basis of primary or secondary data from nine studies, the RR for all types of exposures against acute injury was 0.647 (0.502–0.836, I²=73%, χ² p<0.001). One study had strength training as exposure, two studies did proprioception training and the remaining six studies were from the group of multiple exposure studies. Sensitivity analysis of eight intention-to-treat analysed studies (Soderman et al was excluded) showed a RR 0.615 (0.470–0.803, I²=75%, χ² p<0.001).
Six studies provided data on overuse injuries. RR from these six studies was 0.527 (0.373–0.746, I²=19%, χ² p=0.287). All studies in this analysis, except one proprioception training study, were multiple exposure studies. All analysed studies reported intention-to-treat data (figure 3A,B).

Small-study effect
The Harbord test for the total estimate of all 25 studies showed a highly significant small-study effect test. Exposure and outcome subgroups revealed significant test for only the multiple exposures group. See online supplementary eFigure 7 for modified-Galbraith plot and online supplementary eTable 3 for Harbord tests.

DISCUSSION
An overall RR estimate for physical activity for injury prevention, adjusted for clustering effects, was 0.632 (0.532–0.750), and slightly lower when sensitivity analysed by intention-to-treat (RR 0.607 (0.501–0.735)). A preventive effect of this size should be considered convincing, but the analysis was heterogeneous and the result is, therefore, clinically useless. However, it also suggests that some types of interventions may prove better than others.

Stretching did not show any protective effect (RR=0.961 (0.836–1.106)), while strength training proved highly significant (RR 0.315 (0.207–0.480)). Results from stretching and strength training studies were not heterogeneous despite different programmes were used and outcomes of interest were different. This points towards a strong generalisability of results. Proprioception training and multiple exposure programmes were also effective (RR=0.480 (0.266–0.864) and 0.625 (0.477–0.820), respectively), but results were relatively heterogeneous.

The effect estimate of stretching and proprioception training analyses in this article corresponds to earlier reviews. Our data do not support the use of stretching for injury prevention purposes, neither before or after exercise, however it...
should be noted that this analysis only included two studies on army recruits and one internet-based study on the general population. Strength training showed a trend towards better preventive effect than proprioception training and proved significantly better than multiple exposure studies, even though all multiple exposure studies included a strength training component. Further research of strength training for a wider range of injuries is still needed, as our analyses suggest great sports injury prevention potential for this type of intervention. With a growing number of randomised controlled trials containing numerous exposure types, it was of interest to assess intervention studies with multiple exposures separately, although, as expected, still being a heterogeneous subgroup. Though it makes intuitive sense to design an array of exposures for prevention of all injuries, it is important to note that each component may be reduced quantitatively and/or qualitatively by doing so. Multiple exposure programmes may therefore reduce the proportion of proven beneficial exposures and consequently reduce the overall preventive effect on sports injury. Additionally, the risk of designing too extensive prevention programmes will unavoidably be enhanced with growing amounts of applied exposures and compliance may suffer as a consequence. Although most multiple intervention studies in this analysis were well designed and carried out in a satisfactory way, this subgroup did not exhibit an unambiguous preventive effect on sports injuries. Our findings suggest that designs of multiple exposure interventions should at least be built from well-proven single exposures and that further research into single exposures remains pivotal.

When analyses were stratified by outcome, both acute (RR 0.615 (0.470–0.803)) and overuse (RR 0.527 (0.373–0.746)) injuries were effectively reduced by preventive physical activity, although overuse injuries fared slightly better.

Five of six studies analysing overuse injuries were multiple exposure studies, and estimates were not particularly heterogeneous. Six of nine studies analysing acute injuries were multiple exposure studies with heterogeneous effect sizes. It is not possible to derive which parts of these interventions manifested the preventive effect. Future studies should report acute and overuse injuries separately and test specific exposures against these in order to acquire further knowledge in this import area.

Strengths and limitations
The aim of this meta-analysis was to aggregate a wide array of populations, exposures and outcomes to augment the external validity while maintaining the suitability of combining studies. Physical activity is broadly defined and populations include army recruits, recreational and professional athletes. In this regard, it should be pointed out that the diversity of included studies should not be interchanged with the I² measure of statistical heterogeneity, which exclusively concerns inconsistency in effect sizes. The statistically homogeneous analyses of strength training and stretching studies differing in population, intervention, and outcome, prove the generalisability of results. The statistically heterogeneous analyses of this meta-analysis should be interpreted with caution as this heterogeneity could arise from true variation (diversity in design) and/or artefactual variation (bias by conduct, attrition, etc).

Omission of intention-to-treat analysis and cluster adjustment are two sources of potentially serious bias. As compliance to intervention programmes appears to vary and remains a disputed phenomenon, the analysis by intention-to-treat plays a central role in the robustness of results.⁶⁷-⁶¹ In the present meta-analysis we extracted data from intention-to-treat analyses whenever possible and performed sensitivity analysis by exclusion of five studies with no report of intention-to-treat analysis. Contrary to the expected more conservative effect estimate, the intention-to-treat sensitivity analyses revealed even more beneficial effect estimates. As a result we can conclude that physical activity as primary prevention against sports injuries is effective, even if it has been argued that compliance issues could diminish the implementation and effect of these programmes. We speculate the above to result from an association between using intention-to-treat analysis and study conduct in general. For example, Brushoj et al. added concurrent training in the critical high risk period of military training initiation, which intuitively appears detrimental to overuse injuries. Soderman et al. exhibited several methodological issues and reported adverse effects of major injuries that have not been reproduced by other studies. None of them analysed by intention-to-treat and exclusion of such studies improved the quality of included studies and subsequently the effect estimate.

Cluster adjustment is similarly important in order not to overestimate the power of the study. A strength of this meta-analysis is the adjustment of these studies that report the same effect estimate but underestimate the width of CIs. Corresponding authors of studies without cluster adjustment were contacted and three provided data for p calculation. For the remaining nine studies we calculated an average p value extracted from 12 reported values of 10 studies that performed correct adjustment methods. This caused, in some cases, a dramatic, down-regulation of effective sample size which affected the study weight in the quantitative analyses.

A short discussion of the allocation concealment and participant blinding quality assessments is advocated. As true participant blinding is frequently argued to be impossible in sports injury prevention and allocation, concealment makes less sense in non-pharmacological interventions, these quality assessment items should be interpreted with caution. In spite of this, some of the included studies made qualified efforts to alleviate these, which, in this review, resulted in a lower risk of bias judgement. The domain-based tool was chosen as evaluation tool of this review as recommended by the Cochrane collaboration with the most convincing validation evidence in this area. Although not being perfectly suited for assessment of sports injury prevention studies, assessment of these parameters still holds relevance as these factors can greatly influence analyses.⁶² ⁶³

A Harbord’s small-study effect test and a modified Galbraith’s plot were performed for this meta-analysis to assess publication bias. The small-study effect test for the total estimate was highly significant, while the multiple exposures subgroup was the only subgroup showing a statistically significant test. According to Egger et al.⁶⁴ ⁶⁵ significant small-study effects may arise from a number of reasons, including true publication bias, heterogeneity, chance, and methodological differences between smaller and larger studies. As the p value of the small-study effects increased when the total estimate test was divided into less heterogeneous subgroups, it is likely that a substantial part of the total estimate small-study effect originates in heterogeneity. Owing to the relatively heavy burden of implementing physical activity interventions, it should be noted that smaller studies often would be able to pay greater attention to the intervention for each team/individual, thereby enabling them to obtain more thorough intervention quality. Hence, a methodological difference may exist as well.

CONCLUSION
In general, physical activity was shown to effectively reduce sports injuries. Stretching proved no beneficial effect, whereas...
multiple exposure programmes, proprioception training, and strength training, in that order, showed a tendency towards increasing effect. Strength training reduced sports injuries to less than one-third. We advocate that multiple exposure interventions should be constructed on the basis of well-proven single exposures and that further research into single exposures, particularly strength training, remains crucial. Both acute and overuse injuries could be significantly reduced, overuse injuries by almost a half. Apart from a few outlying studies, consistently favourable estimates were obtained for all injury prevention measures except for stretching.

What this study adds

This meta-analysis provides quantitative effect estimates of different exercise programmes on sports injury prevention. Comparison of exposures reveals a highly effective strength training estimate, significantly better than multicomponent studies.

Acknowledgements The authors would like to thank Thor Einar Andersen, associate professor, Department of Sport Medicine, Norwegian School of Sport Sciences and Ashley Cooper, professor, Centre for Exercise, Nutrition and Health Sciences, University of Bristol for comments and manuscript revision.

Contributors All authors of this paper have contributed substantially to conception and design, analysis and interpretation of the data, drafting the article or revising it critically for important intellectual content and final approval of the version to be published.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES