Journal of Clinical Epidemiology 140 (2021) 111–124

Journal of Clinical Epidemiology

REVIEW

Why methods matter in a meta-analysis: a reappraisal showed inconclusive injury preventive effect of Nordic hamstring exercise

Franco M. Impellizzeri[a,*], Alan McCall[a,b], Maarten van Smeden[c]

[a] University of Technology, Faculty of Health, Sydney, New South Wales, Australia
[b] Arsenal Performance and Research Team, Arsenal Football Club, London, UK
[c] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, the Netherlands

Accepted 7 September 2021; Available online 11 September 2021

Abstract

Objectives: The Nordic hamstring exercise (NHE) has been strongly recommended to reduce hamstring injuries in previous meta-analyses (50% reduction in risk of injury). To underline the importance and impact of adopting appropriate methodology for evidence synthesis, we revisited the study selection, reanalyzed and updated the findings of the most recent meta-analysis.

Study Design and Setting: Only randomized control trials (RCT) using NHE as one of the prevention arms were selected. Summary effects for risk ratios (RR) for original studies included in the earlier meta-analysis, and new studies identified (update), were re-estimated under the random-effects model and presented with 95% confidence intervals (CI) and prediction intervals (PI). Tentative recommendations were provided according to the Grading of Recommendations Assessment, Development and Evaluation.

Results: Only five RCTs out of the 15 studies included in the earlier meta-analysis randomized to NHE. Our update revealed one additional RCT. The point estimate (RR) for the five RCTs previously considered RCTs was 0.56 (95% CI, 0.20–1.52; 95% PI, 0.06–5.14, parametric, and 0.13–1.80, nonparametric). After the update, the RR was 0.59 (95% CI, 0.27–1.29; 95% PI 0.10–3.29, parametric, and 0.17–1.52, nonparametric).

Conclusion: Contrary to the conclusions of a recent meta-analysis, as well as earlier meta-analyses, by using more appropriate methodology, the evidence underpinning the protective effect of NHE so far remains *inconclusive* and mostly derived from high risk of bias RCTs. At best, only conditional recommendation can be provided (for soccer) and future RCTs are warranted.

Keywords: Research synthesis; Injury; Prevention; Heterogeneity; Meta-analysis; Bias; Sports medicine

1. Introduction

Injury prevention strategies have been investigated for decades by international sports organizations and research groups [1]. Among these prevention strategies, the use of the so-called Nordic hamstring exercise (NHE) is widely considered an established and evidence-based prevention method to reduce hamstring muscle injury risk. For example, findings from a recent Delphi survey involving practitioners from top professional European soccer clubs indicated NHE is considered one of the two most effective muscle injury prevention strategies [2]. Accordingly, there are consistent calls in the recent literature for the NHE implementation as preventive strategy and criticism when it is not adopted [3–6]. These calls and recommendations are considered *evidence-based* given the availability of level I evidence, that is evidence coming from meta-analyses:

Conflict of interest: none.

Financial support: none.

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Franco M. Impellizzeri contributed to the conceptualization, methodology, formal analysis, and writing the original draft; Alan McCall contributed to the investigation, and writing (Review & Editing), Maarten van Smeden contributed to the conceptualization, and writing (Review & Editing).

* Corresponding author. F.M. Impellizzeri. Tel.: +61 2 95145204.

E-mail address: Franco.Impellizzeri@uts.edu.au (F.M. Impellizzeri).

https://doi.org/10.1016/j.jclinepi.2021.09.007
0895-4356/

What is new?

Key findings

- We demonstrated the importance of appropriately conducting and interpreting meta-analyses by updating and re-analysing one of the most commonly read meta-analysis used to support that Nordic hamstring exercise prevents sports hamstring injuries.
- We identified and selected five randomised controlled trials, out of the 15 previously used studies in the recent meta-analysis, using Nordic hamstring exercise as one of the experimental arms.
- Summary estimates ranged from highly beneficial to harmful. Considering that the five studies were at high risk of bias and one with some concerns, the effect of Nordic hamstring exercise on prevention of hamstring injuries must still be considered inconclusive.

What this adds to what is known?

- Our study is in conflict with the results of earlier meta-analyses including the same original studies, that have claimed a 50% reduction in hamstring injuries. The evidence underpinning the strong recommendation to implement Nordic hamstring exercise to prevent hamstring injuries is weak and mainly based on high risk of bias studies.

What is the implication, what should change now

- Following GRADE, only conditional recommendations might be provided, and mainly relative to one team sport (soccer). Those implementing Nordic hamstring exercises as preventive strategy should downsize expectations and being aware that expected effects are varying and uncertain, ranging from considerable reduction in the risk of hamstring injury to a substantial increase in injury risk. Those implementing this exercise in other sports should be aware that it is a personal choice not supported by evidence. Future higher quality randomized studies are needed.

two specific on NHE [7,8], one umbrella review including the two previous meta-analyses [9], and one more generic addressing various preventive interventions [10]. All meta-analyses concluded that NHE reduces injuries by 50% or more, and recommended to policy-makers, clinicians and coaches to implement this intervention in light of this line of evidence. The most recent and updated of these meta-analyses has been published in 2019 (online first in 2018) with a strong and unequivocal title: "*Including the Nordic hamstring exercise in injury prevention programmes halves the rate of hamstring injuries: a systematic review and meta-analysis of 8459 athletes*."[7] This article, in only 2 years, has been visualized more than 55 thousand times, >30 thousands full text visualization and downloads, 57 citations and has reached an impressive altimetric score of 478 (https://bjsm.bmj.com/content/53/21/1362.altmetrics; accessed August 5, 2021). This highlights the impact this article has had, and continues to have, within the sports medicine community and among sports practitioners.

Upon reading of these aforementioned meta-analyses, a series of methodological aspects deserve particular consideration, which we aimed to address in the present reanalysis and update:

1) These meta-analyses included studies that used the NHE together with other exercises. However, if the aim was, as stated in previous meta-analyses, to understand whether "Nordic hamstring exercise prevents hamstring injuries *when included* (emphasis added) as part of an injury prevention intervention" [7] or "to investigate the effectiveness of the injury prevention programs that *included* the NH exercise" [8], only studies comparing the effects of programmes with and without the NHE should have been selected. The NHE is the intervention of interest, and the implicit counterfactual of the study aim is: "when [NHE] *is not included* as part of an injury prevention intervention." Combining programmes or other interventions including NHE with other exercises precludes understanding the clinical utility of the NHE itself as an injury prevention strategy. These studies, indeed, implemented and investigated the effect of combining various exercises and not only the NHE (eg, FIFA 11+ prevention program consisting of nine exercises and three progressive levels of difficulty) [11]. Clearly, it is not appropriate to attribute the total average effect of a combination of exercises to NHE alone. The inclusion of combined prevention studies (vs. no prevention) emerged in all previous meta-analyses [7,8,10].
2) Combination of results from different study designs (eg, cohort, nonrandomized and randomized trials), as done in the most recent and updated review,[7] to arrive at a common summary effect is not recommended unless in particular circumstances dictated by the absence of randomized trials, or having randomized trials but where the main research question has been addressed only indirectly or incompletely [12].
3) While the presence of heterogeneity (suggesting real differences in preventive effect across studies) was reported, this was not considered, either computationally or in the interpretation [7,8,10]. The authors, indeed, discussed clinical implications ignoring its relevance for the generalization of findings from a meta-analysis to future situations. Notably, in the presence of heterogeneity, it is recommended to report a 95% prediction interval that reflects a distribution of true mean treatment effects over different settings, like when applying in a new study or in practice [13–16].

4) It is important to not only interpret the point estimate but also the values within the lower and upper bound of the interval estimates that represent effects that are compatible/plausible with the data. This is important for both confidence and prediction intervals, though with the former, this is less informative in the presence of heterogeneity.
5) The previous meta-analyses calculated risk ratios, but some risks were calculated using as numerator the number of injuries including recurrences. An average risk for a group, incidence proportion, is the number of subjects developing an event of interest (eg, injuries) during a time period, divided by the number of subjects followed for the time period [17]. Being a proportion, risks range from 0 to 1, where the maximum value for the numerator corresponds to the total number of subjects followed. Adding recurrent injuries to the numerator is a violation of the computational assumption (unless an appropriate recurrent event analysis is conducted) and hampers the interpretation of risk as a probability [17,18].
6) Injury prevention recommendations should be based on various aspects coming from the published research such as limitations of the original studies, inconsistencies, indirectness, imprecision and publications bias, thereby considering four domains (estimates of effect for desirable and undesirable outcomes of interest, confidence in the estimates of effect, estimates of values and preferences, and resource use) as suggested by The Grading of Recommendations Assessment, Development and Evaluation (GRADE) [19,20].

A recent article has underlined how failure to follow methodological standards and misuse of resources can be misleading and potentially unethical [21]. In the current article, in order to sensitize readers to the importance of applying correct methodology, we showed how both a more appropriate selection of studies and reanalysis accounting from between-study for heterogeneity resulted in different conclusions to claims made in the original meta-analyses. With the aim to re-examine the effectiveness of NHE as preventive strategy for hamstring injuries in athletic populations, we used the same information provided in the most recent meta-analysis while addressing the aforementioned concerns. We also updated the meta-analysis by van Dyk et al. [7], whose search strategy covered until 2018. Finally, the certainty of the evidence and strength of recommendations were also provided according to the GRADE [19].

2. Methods

2.1. Study selection, inclusion and exclusion criteria

Only randomized controlled studies using NHE as one of the experimental arms as a stand-alone treatment were considered for evidence synthesis. Accordingly, we deviated from the inclusion criteria (#2) of the original meta-analysis [7]. The original inclusion criteria, as reported by the authors, were: (1) the population to be any athletes participating in any sporting activity, (2) the intervention to be the NHE [or any programme that included the NHE: criterion eliminated], (3) the comparison to be usual training or other prevention programmes, which did not include the NHE, (4) and the outcome to be the incidence or rate of hamstring injuries. Therefore, the new exclusion criteria were: nonexperimental designs and studies combining NHE with other exercises not performed by the control group. We applied no language restrictions and grey literature was also considered. Since this is a reanalysis of a previous meta-analysis with stricter inclusion criteria, we used the list of studies and data presented by van Dyk et al. [7] following additional scrutiny of the information reported (Tables 1 and 2). In the context of the current reanalysis, we used the conventional term effectiveness (effect under "real world," nonideal conditions) [22,23], instead of efficacy (effect under more controlled conditions) [22,23], because the RCTs were all identified as pragmatic trials, given the lack of blinding, heterogeneous samples, with interventions delivered under real setting and we focused on the effect of being assigned to the intervention.

2.2. Update and data extraction

We updated the original meta-analysis by using the same search strategy reported in the supplementary file of the original meta-analysis [(Nordic OR Russian) AND hamstring AND injur*] and in the registration. The update was conducted using the modified inclusion and exclusion criteria presented in the previous section. The search was conducted from August 2018 (final search date of the original study) until May 2021, on four databases: PubMed, CINAHL, Web of Science and OpenGrey (further details of search strategy and results in the Supplementary file #1). Study selection was conducted independently by two investigators (F.M.I. and A.M.C.) and uploaded in Endnote, following a multi-stage process in which potentially eligible studies were first identified from screening titles and abstracts and then assessed through full text review. No automatic tools were used. Data extracted and tabulated for the update were: authors, location, samples size and number of hamstring injuries for each group (primary outcome), sex, age, level of competition, design, protocol of the intervention, sport. Data were extracted by one research and checked by a second author. Extracted data were compared with any discrepancies resolved through discussion. Discrepancies with the data extracted in the previous meta-analysis, were resolved trying to contact the author of the original studies. This systematic review with meta-analysis was reported according to PRISMA 2020 [24]. A checklist showing how each

Table 1. Characteristics of the original studies included in van Dyk et al. (2019) and the additional study from the update

Study(ID Authors)	Sex	Age (yr)[a]	Design	Level	Sport	Intervention	Control	Includedin the update	Reasons for exclusion
From the meta-analysis by van Dyk et al. (2019)									
#1 Gabbe et al. 2006 [36]	Male	17.4–36 (18–35)	RCT	Australian Amateur	Australian Football	NHE	Basic stretching	√	
#2 Brooks et al. 2006 [38]	Male	19–34 (18–40)	Cohort prospective	English Premiership	Rugby Union	Strengthening, Stretching and NHE	Strengthening (control 1), Strengthening and Stretching (control 2)		Design
#3 Arnason et al. 2008 [39]	Male	Not reported	Unclear (non-randomized parallel groups and interrupted time series)	Norwegian and Iceland Elite League and First Division	Football (soccer)	Season 2001 (Iceland) warm up, flexibility and NHE; Season 2002 (Iceland & Norway) warm-up and strength.	Season 1999 (Iceland) and 2000 (Iceland & Norway) usual training; Season 2001 (Norway) warm-up and flexibility.		Design
#4 Soligard et al.2008	Female	13–17	Cluster RCT	Norwegian youth (clubs)	Football (soccer)	F-MARC/FIFA 11+ (multicom-ponent)[b]	Usual warm up		NHE not in-dependent variable
#5 Enge-bretsen et al. 2008 [37]	Male	Not reported (17–35)	RCT (nested in a cohort study)	Norwegian 1st, 2nd, or the top of the 3rd division	Football (soccer)	NHE	Usual training	√	
#6 Petersen et al. 2011 [35]	Male	23.0 (s = 4.0) and 23.5 (s = 4.0) (18–40)	Cluster RCT	Top 5 Danish divisions (profession-als to amateurs)	Football (soccer)	NHE	Usual training	√	
#7 Owen et al. 2013	Male	28.6 (s = 3.75) (18–40)	Interrupted time series	Scottish Premier League (profession-als)	Football (soccer)	Multicomponent prevention intervention (12 exercises including NHE)	Usual training		Design and NHE not in-dependent variable
#8 Grooms et al. 2013	Male	18–25	Nonrandomized controlled trial	NCAA Division III collegiate players	Football (soccer)	F-MARC/FIFA 11+ (multicom-ponent)[b]	Standard NSCA dynamic warmup		Design and NHE not in-dependent variable
#9 Sebelien et al. 2014 [33]	Male	18–36 (18–39)	RCT	Adult level three and four Norwegian semiprofes-sional	Football (soccer)	NHE & passive hamstring stretching	Usual training and general stretching (not specified whether general stretching include hamstrings)	(√)	
#10 Seagrave et al. 2014 [40]	Male	26.6 (s = 2.7) 22.8 (s = 3.1) (18–40)	Nonrandomized controlled trial	Major and minor league teams	Baseball	NHE	Control 1, usual training; control 2, noncompliant (<3.5 mean reps/wk)		Design

(continued on next page)

Table 1 (*continued*)

Study(ID Authors)	Sex	Age (yr)[a]	Design	Level	Sport	Intervention	Control	Includedin the update	Reasons for exclusion
#11 van der Horst et al. 2015 [34]	Male	24.5 (s = 3.8) (18–40)	Cluster RCT	Netherlands high–level amateurs	Football (soccer)	NHE	Usual training	√	
#12 Silvers-Granelli et al. 2015	Male	18–25	Cluster RCT	NCAA Division I and Division II players	Football (soccer)	F-MARC/FIFA 11+ (multicomponent)[b]	Usual training		NHE not independent variable
#13 del Ama Espinosa et al. 2015	Female	23 (s = 4) team 1 19 (s = 2) team 2 22 (s = 3) team 3	RCT	First and second Spanish division	Football (soccer)	NHE & eccentric band exercise	Frontal leg swings, side leg swings and multiple jumping.		NHE not independent variable
#14 Nouni-Garcia et al.2017	Male	24.7 (s = 3.9) (18–40)	Unclear (maybe case-control nested in a cohort)	Spanish amateurs, first regional division	Football (soccer)	F-MARC/FIFA 11+ (multicomponent)[b]	Usual training		Design and NHE not independent variable
#15 Gonzalez et al. 2017	Male	18.6 (s = 0.1) (18)	Interrupted time series	Young players of a professional Spanish team	Football (soccer)	Multicomponent prevention intervention (8 exercises including NHE)	Usual training		Design and NHE not independent variable
Update 2021									
#16 Hasebe et al. 2020 [27]	Male	15–18	Cluster RCT	Japanese high school players	Football (soccer)	NHE	Usual training	√	

Abbreviation: RCT, randomized controlled trial.

[a] Intervention and control combined or separated by teams/groups, and in parentheses the values reported by van Dyk et al. 2019.

[b] F-MARC/FIFA 11+ is a warmup preventive program including running exercises, plyometric, plank, balance, jumping, squatting and NHE (three levels of increasing difficulty).

PRISMA item was addressed is included in the Supplementary file #2.

2.3. Risk of bias assessment and certainty of evidence grading

We reassessed the risk of bias (RoB) of all the studies using the revised Cochrane “Risk of bias” tool for randomized and cluster trials (RoB 2) using the online resources (https://www.riskofbias.info/welcome/rob-2-0-tool/current-version-of-rob-2) [25]. The GRADE was used to appraise the certainty of the evidence for the studies deemed eligible for quantitative synthesis [19]. Rating and grading were completed by two investigators recording the supporting information and justifications for judgements (presented in the Supplementary file #3, using the most recent templates, 2019 and 2021). Any discrepancies in judgements were resolved by discussion to reach consensus between the two reviewers (FMI and AMC), with a third review author (MvS) acting as referee if necessary.

2.4. Statistical analyses

Evidence synthesis involved analyses of the original and updated data extractions, with outcome measures of interest expressed as dichotomous (injured/non injured). For the analyses of the original data extraction, as in the previous meta-analysis [7], the weighted raw point estimate for the Risk Ratios (RR) was calculated using a random-effects model, with DerSimonian and Laird estimator for the t-statistic, and Mantel-Haenszel estimator used in cal-

Table 2. Details of quasi-experimental (nonrandomized), observational studies and comparison with injury data reported in van Dyk et al. (2019)

Studies	Groups	Checked data		van Dyk et al. 2019		Inconsistencies(in van Dyk et al.)
		Injuries	Sample	Injuries	Sample	
Brooks et al. 2006[a]	Strength Strength, stretch Strength, stretch, NHE	53 32 27 Tot = 112	148 144 200 Tot = 492	CON = 10 NHE = 0 Tot = 10	292 200 Tot = 492	Error in total injury (102 missed); all injury in NHE missed.
Arnason et al. 2008[a]	1999 Iceland, no interventions 2000 Iceland, no interventions 2000 Norway, no interventions 2001 Iceland, warmup, flex, NHE 2001 Norway, warmup, flex 2002 Iceland, warmup, NHE 2002 Norway, warmup, NHE	30 33 34 26 33 7 20 Tot = 183	306 271 252 288 250 180 252 Tot = 1799	CON = 96 NHE = 81 Tot = 177	1218 882 Tot = 2100	NHE groups in original study had 53 (26 + 7 + 20) injuries and not 81; total injuries wrong (177 vs. 183); total sample wrong (1799 vs. 2100)
Seagrave et al. 2014	Noncompliant with NHE Control NHE compliant	7 3 0 Tot = 10	178 34 65 Tot = 277	CON = 10 NHE = 0 Tot = 10	34 65 Tot = 99	Injuries in CON wrong; noncompliant excluded (n = 178) but injuries added to CON
Elerian et al. 2019	Team previous season NHE once a week NHE twice a week	20 (7 recur) 1 (0 recur) 4 (1 recur) Tot = 25	35 17 17 Tot = 69	NA	NA	NA

[a] Data provided by the authors.

culation of Q-statistic plus continuity correction of 0.5 in studies with zero cell frequencies [26]. For the analyses of the updated data extraction, including a recent RCT [27], we estimated the mean summary effect using a random-effects model with Hartung-Knapp adjustment and Restricted Maximum-Likelihood Estimator (REML) estimator for the t-statistic relevant to small-to-medium scale meta-analyses with rare events [28–30]. We used the generalized Q-statistic method to describe the uncertainty around the mean τ-statistic value. The uncertainty around point estimates were presented as 95% confidence interval (CI) [28].

We derived a 95% prediction interval (PI) to describe the range of true effects expected in 95% of future, similar studies in the presence of between-study heterogeneity using parametric and nonparametric methods [31,32].

Given the anticipated low number of studies, no subgroup or sensitivity analysis was planned. Results were presented in Forest plots. For the readers information and educational purposes only (ie, not specifically relevant to the current meta-analysis), we have also presented the summary effects calculated under the fixed effect model (and additional explanations of the different models in Appendix). The codes of the analyses using *meta* package in R (version 4.0.5, R Foundation for Statistical Computing, Vienna) are available as Markdown file in Supplementary file #4.

3. Results

3.1. Study selection for the reanalysis

As shown in the PRISMA 2020 flow chart (Supplementary file #5) and Table 1, only eight studies out of the 15 studies included in the review by van Dyk et al [7]. did use NHE as a stand-alone treatment. Among these, five were RCTs [33–37] and three [37–40] used nonrandomized or observational designs. Checking the information presented by the previous authors (for the eight studies using NHE as independent variable only) inconsistencies emerged, some of which were minor (different age ranges and different design classification) and some actual errors such reversed sample sizes (n for control inverted with n for intervention)

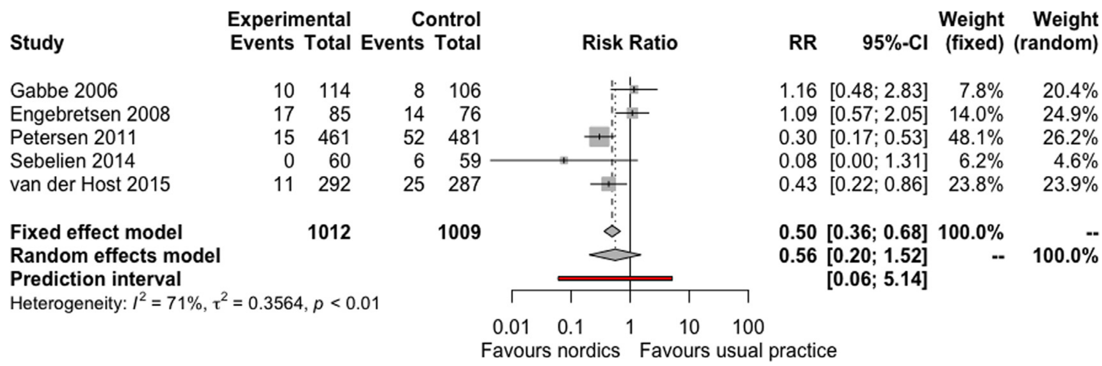

Fig. 1. Forest plot of the five randomized controlled trials selected from the meta-analysis by van Dyk et al [7]. using the Nordic hamstring exercise as independent variable (nonparametric 95% prediction interval, 0.13–1.80).

[33], wrong number of injuries [40], number of injuries not coherent with the original paper [38,39], and not reported in the original paper [39].

Five authors were contacted for additional information [10,33,34,38,39], and four answered and provided explanations or the data not reported in the papers [10,34,38,39]. Most of the missing information refer to the observational studies that we did not use to calculate the summary estimate in the current meta-analysis (Table 2). Only van der Horst et al [34]. was contacted to ask clarification about the preplanned statistical analysis [41]. For the RCT included in the current reanalysis and update, the error (sample size reversal) for the RCT by Sebelien et al [33]. was corrected and for Engebretsen et al [37]. we used the number of injuries excluding recurrent event instead of the combined figures reported in the meta-analyses of Al Attar et al [8]. and van Dyk et al [7].

3.2. Reanalysis

When the appropriate RCTs using NHE as the treatment variable were considered and recurrent injuries excluded, the summary estimates (ie, RR) of the five RCTs included in the meta-analysis by van Dyk et al [7]. and calculated using the same methods ranged from 0.20 to 1.52 (point estimate 0.56). Furthermore, the width of the 95% PI suggested the effect could fall anywhere in the range of 0.06–5.14 in future similar studies or implementations (Fig. 1). When a nonparametric method was used, the 95% PI ranged from 0.13 to 1.80. In the original meta-analysis, the RR of the 15 studies was 0.49 (95% CI, 0.32–0.74).

3.3. Update of the meta-analysis

For the update, we identified 210 eligible records (see Supplementary file #5). After duplicates removal, 120 records were screened and full text was retrieved for 2 records, and one was excluded since it did not meet the inclusion criteria (not an RCT) [42]. A further search was conducted using the Connected Papers online platform (https://www.connectedpapers.com/) but no further articles were found.

When this recent RCT was added and a more appropriate analysis conducted accounting for the small number of studies, the 95% CI of summary estimates (RR) ranged from 0.27 to 1.29 and 95% PI ranged from 0.10 to 3.29 and from 0.17 to 1.52 using parametric and nonparametric method, respectively (Fig. 2). The mean τ-statistic value was 0.54 (95% CI, 0.12–2.30). That is, when updated with a recent study, the heterogeneity remained substantial.

3.4. Risk of bias and quality of evidence

The traffic lights plots of the risk of bias assessment according to the RoB 2 for randomized and cluster randomized studies are present in Figures 3A and B, respectively. The motivations for the rating in each domain for each study is presented in the Supplementary file #3 to allow the reader or previous authors to understand the reasons leading to the rating.

The evidence profile according to the GRADE with the rating explanations is presented in Table 3. The resulting quality of evidence was low.

4. Discussion

Recent meta-analyses concluded that NHE decreases the risk of injuries by 50% or more and, as a consequence, this prevention strategy is recommended and more efforts to increase its implementation are strongly encouraged [3–6,8,10]. However, we identified several methodological concerns in previous meta-analyses that, when properly addressed, have an important effect on the meta-analytic summary estimates. Our study showed that the effectiveness of NHE as an injury prevention strategy to reduce injury risk remains uncertain.

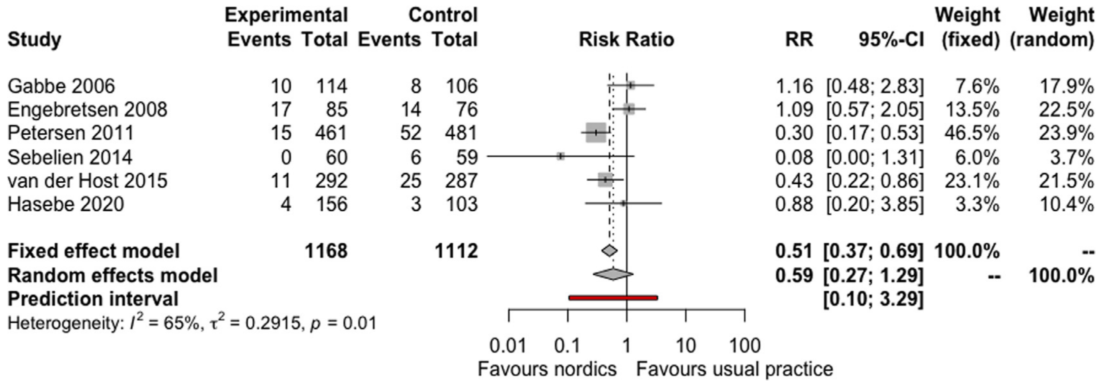

Fig. 2. Forest plot of the selected randomized controlled trials plus a recent additional study (Hasebe et al., 2020) using the Nordic hamstring exercise as independent variable (nonparametric 95% prediction interval, 0.17–1.52).

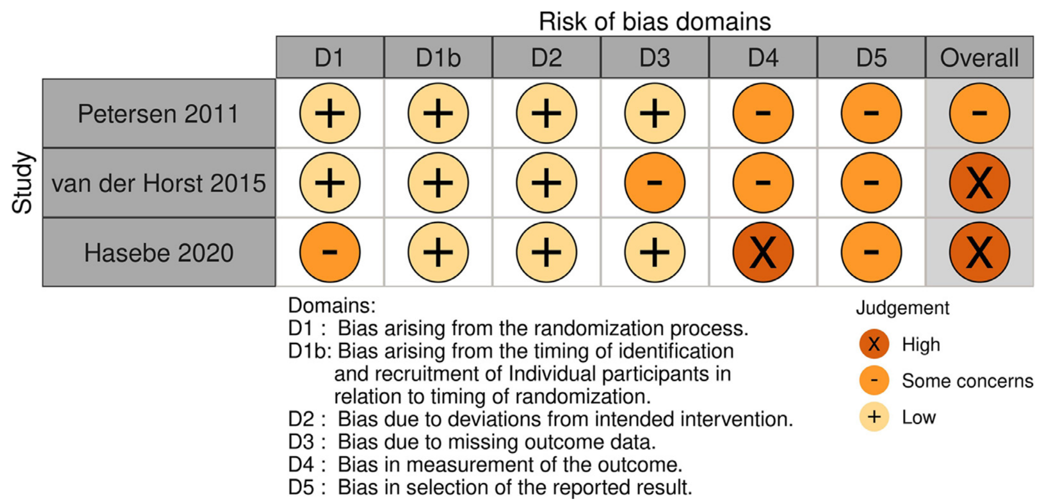

Fig. 3. Traffic lights plot presenting the risk of bias for cluster-randomized (A) and randomized (B) trials.

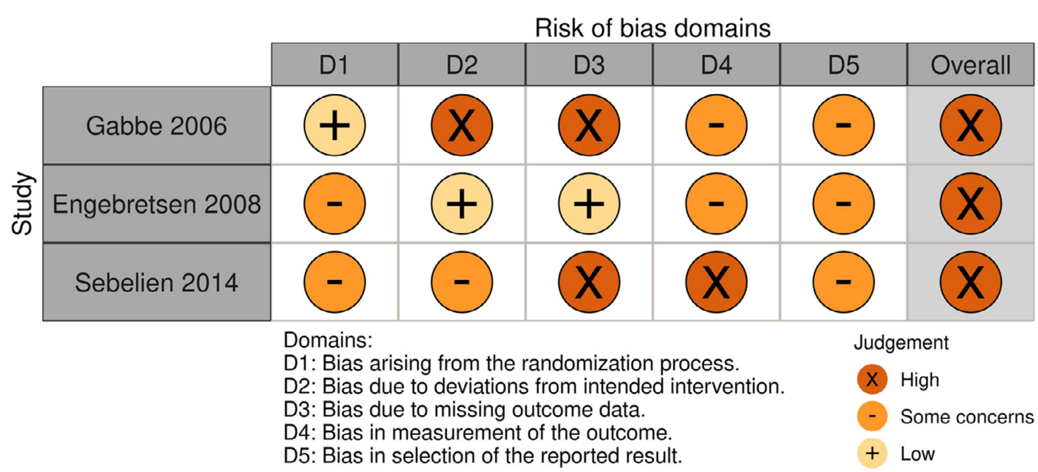

Fig. 3. Continued

4.1. Selection of studies based on the proper independent variable and research design

To isolate and investigate the effect of an intervention, it is well known that researchers should manipulate the intervention by adding the intervention to one arm only, and compare the effects in two otherwise exchangeable groups [43]. If the investigated intervention is NHE, this necessitates one group completing the NHE and one group completing the usual training. However, if the researcher

Table 3. Evidence profile according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE)

	Limitations	Inconsistencies	Indirectness	Imprecision	Publication bias	Quality of evidence
Judgement	Serious limitations	Serious limitations	No serious limitations	Serious limitations	Unclear	⊕⊕◯◯ Low
Explanation	Evidence coming from studies with high risk of bias or some concerns in various domains according to the Cochrane RoB 2.	High heterogeneity Wide variance of point estimates across studies	However, most studies (five out of six) on soccer; application in other sports may be limited by indirectness.	Both 95% CI and PI range from protective effect to harmful	Few studies for formal testing and interpretation difficult.	Limited confidence in the effect estimates
Downgrade	(-1)	(-1)	(No)	(-1)	NA	

implements NHE *and* other exercises, and these additional exercises are not also added to the control group, it is clearly not possible to isolate the effect of NHE intervention under investigation. As a consequence, any effect cannot be attributed to NHE but rather to the combination of exercises, which is the real treatment variable manipulated in the design. The use of ambiguous wording such as "*NHE alone or in combination with an injury prevention program*" [8] or "*when included as part of an injury prevention intervention*" [7] does not solve this problem because, to examine causal treatment effects, a control group performing the exercises of the prevention program but not the NHE would be anyway needed. We concede, however, in a previous meta-analysis authors have acknowledged this as a limitation [8], but still concluded their meta-analysis was the first "*to report the efficacy of the NH exercise in preventing hamstring injuries among soccer players*" and that their findings "*can assist policy makers involved in decisions about the implementation of measures to reduce soccer injuries*" [8]. Acknowledging this does not make the summary estimate right, especially if not considered in the conclusions.

Furthermore, combining all designs as in the meta-analysis of van Dyk et al. [7], such as merging RCT, quasi-experimental (nonrandomized trials) and observational (cohort studies) is not recommended [12]. Nonrandomized studies and even more so observational studies based on spontaneous participation are subjected to biases that are different than RCTs, and indeed are usually assessed using different RoB tools. The Cochrane [12] suggests two main reasons for which nonrandomized can be "acceptable" to examine the effect of an intervention. First, the lack of RCT addressing the research question directly and, second, when the research question itself cannot be addressed with an RCT design. However, none of these are reasons applicable to NHE. Even when the use of nonrandomized studies is acceptable, only strong nonrandomized studies should be included [12], that is studies accounting for the various sources of biases, confounding and preferably preregistered. The selected three observational studies plus the one located in the update and using NHE as independent variable are unlikely to be "strong." With the exclusion of one more typical cohort study [38], the other three used unusual designs combining sort of case-crossover/interrupted time series nested in a nonrandomized [39] or randomized trial [42]. Importantly, none of these studies tried to even consider any confounders, some excluded participants or recategorized them after assigning the intervention, and all were not preregistered. Combining all of these studies with different designs, most of which were poor, and mixing the independent variables as done in the meta-analysis by van Dyk et al [7] is clearly misleading. Even more so if data in the meta-analysis does not match the figures of the original papers (Table 2).

In the presence of heterogeneity, the 95% CI is less informative especially for decision-making. However, for comparison with the previous meta-analysis, we briefly discuss the CIs of the current re-analysis. By selecting appropriate studies (according to design and treatment variable) taken from the same original study list provided by van Dyk et al [7], the cumulative evidence underpinning the protective effects of NHE is *inconclusive* with 95% CI including protective (RR = 0.20) but also harmful effects (RR = 1.52). This uncertainty is also larger than the one reported in the original secondary analysis using only eight RCTs (RR, 0.32–0.85). However, also in this secondary analysis, the estimates of van Dyk et al [7]. included RCTs using NHE in isolation and in combination with other exercises. Furthermore, the authors included among RCTs a nonrandomized study [40], while a proper RCT [33], included in our reanalysis, was excluded. Updating and including an additional recent RCT, the true intervention effects across studies are compatible with effects ranging from 0.27 to 1.29, that is, compatible with both very protective but also increased injury risk.

4.2. Taking into account the between-study heterogeneity

When a researcher aims to provide clinical implications, that is, what preventive effect to expect on the injury risk by implementing the NHE, the 95% PI should be considered. Importantly, a 95% PI describes a range of effects we may expect in 95% of future similar studies involving random samples of patients or, in this context, athletes on which we would like to decrease the hamstring injury risk by implementing the NHE [13,14]. The 95% PI is calculated taking into account the between-study heterogeneity. As expected, using the five selected RCTs from the previous meta-analysis [7], the 95% PI resulted wider than the 95% CI, with the lower and upper limits in our reanalysis ranging from 0.06 to 5.19 (parametric) and from 0.13 to 1.80 (nonparametric). Likewise, the 95% PI for the update meta-analysis ranged from 0.10 to 3.29 (parametric) and from 0.17 to 1.52 (nonparametric), that is again compatible with very protective but even harmful effects, whatever the method of calculation. Prediction intervals and predictive distributions may be less robust to violations of the normality [31,44] and when a small number of studies are available, this can influence the width of the interval. Although, technically, PI can be calculated when more than three studies are available [14], the Cochrane encourage the use of PI when more than 10 studies are available and there is no asymmetry in the funnel plot [45]. InHout et al [13]. have shown acceptable estimates with more than six studies. Nevertheless, we have addressed this issue by also calculating the PI using nonparametric methods (more conservative), which does not assume that the underlying heterogeneity follows a normal distribution [31]. None of the previous meta-analyses have provided the PIs or explained why they were not calculated, despite recommending implementation that implicitly referred to expected effects in future applications. If PI was not calculated because of the small sample of studies, first, this would deserve discussion and, second, this would anyway limit the strength of the recommendations. The issue of heterogeneity and its implication in terms of interpretation and computationally was not addressed in any of the previous meta-analyses. Furthermore, it should be reminded that random effect models do not address or explain heterogeneity but only take this into account, statistically, in the calculation of the summary estimates [14]. The potential sources of heterogeneity must be investigated. If the limited number of studies does not allow exploration of the source of heterogeneity using for example (preplanned) sensitivity analyses or meta-regression [46], this limitation should be considered in the interpretation of the results, including how they should (or not) be used for decision-making.

4.3. Interpretation of the summary estimate based on the underlying model and uncertainty

Stressing and focusing on the point estimate as done in the previous meta-analyses is in contrast with the assumption of the random-effect model used in all the meta-analyses [7,8,10]. Indeed, claiming that NHE reduces injuries by 50% is like to interpret the point estimate as in a fixed-effect model, that is, using this as an estimation of the shared common true effect. Instead, the point estimate in a random-effect model is an estimation of the mean of the *distribution of true effects* (see Appendix). While this seems to be a frequent error also in epidemiology as shown by Riley et al. [47], researchers should interpret the summary results as an estimate of the average effect rather than the common effect, when a random effect model is used [14]. By stressing in all meta-analyses that the NHE should be implemented because doing so halves the risk of injuries, the authors encouraged practitioners to think that 50% is the expected (true) effect. This is a conceptually misleading both because it misinterprets the point estimate generated using the random effect model, and because the *expected* range of effects is given by the 95% PI, as explained in the previous section. Furthermore, the range of compatible values given by the confidence and prediction intervals cannot be ignored. As suggested by Borenstein et al [48]. when there is a substantial variation (wide CIs), it is “more informative in general to shift our focus away from the mean and toward the dispersion itself.”

4.4. Risk of bias of the studies

Overall, the studies were all at high risk of bias given “some concerns” or “high risk” in the various domains. Detailed explanations, criteria and comments of the rating are provided in the Supplementary file #3. The cluster RCTs showed less domain at high risk of bias than the studies randomizing individual participants. When compared with the previous meta-analyses some discrepancies emerged. However, van Dyk et al [7] used the old version of the Cochrane RoB for rating both RCT and observational. Al Attar et al [8]. used an adapted version of the Cochrane RoB for spinal disorders [49]. Both did not report the justifications of the ratings and therefore it is not possible to understand the reasons of the variability and discrepancies in ratings. Clearly, some sources of bias are inevitable such as the difficulty to blind participants to an intervention like NHE. Nevertheless, they remain sources of potential biases. Finally, our reanalysis noted that none of the cluster RCTs analyzed the data conditional to cluster, which would be required, and considered/addressed competing risks (eg, other injuries).

4.4. Recommendations for NHE according to the GRADE

Although all previous meta-analyses provided recommendations, none of them used established methods such as the GRADE that, by providing clear guidelines, may limit biased interpretations. The rating of the quality of evidence (certainty of evidence or confidence in the estimates) for the NHE according to GRADE is low. To evaluate the publication bias, we generated contour-enhanced meta-analysis funnel plot, but the few studies did not allow formal testing of asymmetry, and distribution was difficult to interpret. We therefore did not consider the publication bias domain (unclear), but we added in the Supplementary file #6 the contour-enhanced funnel plot for the readers. While the confidence in the estimates is not sufficiently strong to recommend the use of NHE to prevent hamstring injuries, as instead is consistently reported in the literature and previous meta-analyses, there is also no evidence to recommend against its use.

In these situations, to decide whether to recommend or not an intervention, the GRADE suggests various aspects to consider: (1) quality of evidence, (2) the balance of desirable and undesirable consequences, (3) values and preferences of those affected, and (4) resource use. As shown in Table 1 the overall quality of evidence is low. Hamstring strains and associated burden is one of the most common muscle injuries in sport [50–53]. Decreasing the hamstring injury is certainly of interest but it is also crucial to avoid any situation that can increase the risk (and burden). Since it cannot be excluded that NHE may increase the risk given the 95% CI and PI calculated in this meta-analysis, the balance between the risk and benefit is difficult to evaluate. Deciding whether to implement NHE accepting the risk is more likely something that individual staff should evaluate based on an honest and unbiased presentation of the summary estimates. Values and preferences of athletes is something rarely considered and investigated in the context of prevention programs in general and for NHE in particular. This would be helpful to understand the balance between desirable and undesirable effects by knowing the weight (importance) of the risk and benefit from the athlete perspective (values and preferences). In terms of resources, NHE is certainly a feasible exercise that does not require equipment.

All considered, following the suggestions of GRADE, NHE might be "conditionally recommended" (ie, conditional to athlete, technical and clinical staff preferences, values and context) given the low cost. Alternatively, NHE can be "discretionary recommended," another formulation of the GRADE, based on the opinion of the practitioners and athletes. Nevertheless, it is also important to consider that these (weak) recommendations might be only relevant for soccer players, given that no RCTs have been conducted in other sports, with the exception of one study (with high risk of bias) on Australian Rules Football players. Therefore, in all other sports NHE can be recommended mainly for research that, according to GRADE, is appropriate when (1) there is insufficient evidence to support a decision for or against an intervention, (2) further research can reduce uncertainty about the effects of the intervention, and (3) further research is thought to be of good value for the anticipated costs (that for NHE implementation are low and hence it certainly is a scalable intervention).

4.5. Limitations

We have assumed that the original studies have adopted an intention to treat (ITT) approach, but only three studies [34,36,37] clearly referred to ITT analysis. However, from the participants flow chart of the studies and the setting we considered plausible that also the other studies used ITT. Nevertheless, no details about the exact ITT approach were reported in any of the six RCT. Another limitation is the relatively low number (five and six) of studies that were included in the analysis. Although for the update of the meta-analysis we used methods to partially account for the small sample (ie, Hartung-Knapp adjustment and REML estimator for the τ-statistic) [29,30] and the prediction intervals were also calculated using a nonparametric method, the relatively low number of studies may have influenced the 95% CIs and PIs [28]. Whether such a sample of studies is enough to provide informative summary estimates that can be used to provide recommendations certainly deserves consideration. Nevertheless, three meta-analyses have been published and as long as the researchers decide that a meta-analysis is worthwhile, this should be conducted in a rigorous way with results properly presented and interpreted. The limitations such as the small sample should be acknowledged and accounted in the analysis using appropriate methods. These limitations should then be considered when providing recommendations. Finally, we applied the GRADE to suggest what we believe, based on available evidence, are more reasonable and balanced recommendations. However, these were mainly tentative and an example to show the process that should generate recommendations, which clearly would necessitate a more comprehensive and extended panel of experts.

5. Conclusions

The current reanalysis demonstrated the lack of empirical (and conceptual) support to recommendations based on the conclusions that NHE decreases hamstring injuries by 50% and that a large amount of evidence supports the use of the NHE to prevent hamstring injures [4,5,7,8,10,54]. Previous conclusions were unintentionally misinformed by: (1) errors in the study selection criteria, (2) focus on point estimates ignoring the diversity of the effects of the true mean distribution given the underlying model/assumptions, (3) lack of consideration of between-study heterogeneity, thereby failing to estimate the effects that can be expected in future similar settings, and (4) development of recommendations based on personal and potentially biased interpretations without following established and well-defined guidelines. While we hope the previous claims will be confirmed in the future, by now the resulting recommendations based on a more rigorous meta-analysis and the application of the GRADE are, at best, conditional and limited to one sport (soccer). Extension to other sports necessitates further studies.

While the current critical report is directed at sports medicine using a specific sports injury prevention example, the applicability and take-home message are shared with the wider medical community. This is evidenced through a very recent commentary entitled: *"Methodology over metrics: current scientific standards are a disservice to patients and society*" [21]. As suggested by van Calster and colleagues [21], to provide value to patients and society, there is the need to improve the transparency and methodological quality of studies (the backbone of science). Accordingly, with our reanalysis, we have highlighted the importance of using appropriate methods when conducting a meta-analysis and the impact this can have on the conclusions and recommendations that are subsequently disseminated to practitioners. Researchers should take time and care when conducting such a study because findings and conclusions drawn from a published "meta-analysis" have considerable impact and if findings are misleading or misinterpreted, this can create empirically unsupported expectations.

Acknowledgments

We would like to thank the authors of previous original studies and meta-analysis that kindly provided us all the additional information and clarifications: Árni Árnason, John H.M. Brooks, Nick van der Horst, Rok Vatovec and Nejc Šarabon. We also thank Aaron J. Coutts for the comments provided regarding a preliminary version of this manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jclinepi.2021.09.007.

Appendix. A fixed-effect model assumes there is a distribution of common effects (the studies share a common true effect) [48], and the observed mean of each study is given by

$$Y_i = \theta + \varepsilon_i,$$

where Y_i is the observed mean for study i, θ is the common true effect and ε_i is the estimation error, that is the only source of variation $(\varepsilon_i,)$. Instead, a random-effect model assumes a distribution of true effect sizes, where the goal is to estimate the mean of this distribution [48]. In a random-effect model the observed mean (Y) of each study i is indeed given by:

$$Y_i = \mu + \xi_i + \varepsilon_i,$$

where μ is the grand mean, ξ_i is the difference between the grand mean and the true mean for study i (θ_i), and ε_i is the difference between the true mean θ_i and the observed mean Y_i for study i. Here there are two sources of variations, ξ_i andε_i

References

[1] Vriend I, Gouttebarge V, Finch CF, van Mechelen W, Verhagen E. Intervention strategies used in sport injury prevention studies: a systematic review identifying studies applying the Haddon Matrix. Sports Med 2017;47(10):2027–43.
[2] McCall A, Pruna R, Van der Horst N, Dupont G, Buchheit M, Coutts AJ, et al. Exercise-based strategies to prevent muscle injury in male elite footballers: an expert-led Delphi Survey of 21 practitioners belonging to 18 teams from the big-5 European leagues. Sports Med 2020;50(9):1667–81.
[3] Al Attar WSA, Komir R, Alzubeadi A, Bukhari I, Ghulam H. Limited implementation of the Nordic hamstring exercise in professional and semi-professional soccer. J Orthop Trauma Rehabil 2021;28:1–7.
[4] O'Connor S, Lacey P. Can we improve coaches' injury prevention views and implementation practices in the community female Gaelic sport of camogie? BMJ Open Sport Exerc Med 2020;6(1):e000732.
[5] van der Horst N, van de Hoef S, van Otterloo P, Klein M, Brink M, Backx F. Effective but not adhered to: how can we improve adherence to evidence-based hamstring injury prevention in amateur football? Clin J Sport Med 2021;31(1):42–8.
[6] Chesterton P, Tears C. The uptake of the Nordic hamstring exercise programme as an injury prevention strategy in professional cricket in the United Kingdom and barriers to implementation. Phys Ther Sport 2021;50:1–6.
[7] van Dyk N, Behan FP, Whiteley R. Including the Nordic hamstring exercise in injury prevention programmes halves the rate of hamstring injuries: a systematic review and meta-analysis of 8459 athletes. Br J Sports Med 2019;53(21):1362–70.
[8] Al Attar WSA, Soomro N, Sinclair PJ, Pappas E, Sanders RH. Effect of injury prevention programs that include the Nordic hamstring exercise on hamstring injury rates in soccer players: a systematic review and meta-analysis. Sports Med 2017;47(5):907–16.
[9] Raya-GonzÁlez J, Castillo D, Clemente FM. Injury prevention of hamstring injuries through exercise interventions: an umbrella review. J Sports Med Phys Fitness 2021 Online ahead of print. doi:10.23736/S0022-4707.21.11670-6.
[10] Vatovec R, Kozinc Z, Sarabon N. Exercise interventions to prevent hamstring injuries in athletes: a systematic review and meta-analysis. Eur J Sport Sci 2020;20(7):992–1004.
[11] Silvers-Granelli H, Mandelbaum B, Adeniji O, Insler S, Bizzini M, Pohlig R, et al. Efficacy of the FIFA 11+ Injury Prevention Program in the collegiate male soccer player. Am J Sports Med 2015;43(11):2628–37.
[12] Reeves BC, Deeks JJ, Higgins JPT, Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, et al. Chapter 24. Including non-randomized studies on intervention effects. Cochrane handbook for systematic reviews of interventions, 6, Cochrane: version; 2021. updated February 2021.

[13] IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 2016;6(7):e010247.

[14] Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. BMJ 2011;342:d549.

[15] Graham PL, Moran JL. Robust meta-analytic conclusions mandate the provision of prediction intervals in meta-analysis summaries. J Clin Epidemiol 2012;65(5):503–10.

[16] Kontopantelis E, Springate DA, Reeves D. A re-analysis of the Cochrane Library data: the dangers of unobserved heterogeneity in meta-analyses. PLoS One 2013;8(7):e69930.

[17] Rothman KJ. Epidemiology: an introduction. Oxford University Press; 2012.

[18] Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Lippincott Williams & Wilkins; 2008.

[19] Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. ed The GRADE Working Group (Available from guidelinedevelopment.org/handbook, 14 May 2021), 2013.

[20] Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol 2013;66(7):726–35.

[21] Van Calster B, Wynants L, Riley RD, van Smeden M, Collins GS. Methodology over metrics: current scientific standards are a disservice to patients and society. J Clin Epidemiol 2021 Online ahead of print. doi:10.1016/j.jclinepi.2021.05.018.

[22] Haynes B. Can it work? Does it work? Is it worth it? The testing of healthcare interventions is evolving 1999;319(7211):652-653.

[23] Godwin M, Ruhland L, Casson I, MacDonald S, Delva D, Birtwhistle R, et al. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Med Res Methodol 2003;3:28.

[24] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. J Clin Epidemiol 2021;18(3):1–16.

[25] Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019;366:l4898.

[26] Efthimiou O. Practical guide to the meta-analysis of rare events. Evid Based Ment Health 2018;21(2):72–6.

[27] Hasebe Y, Akasaka K, Otsudo T, Tachibana Y, Hall T, Yamamoto M. Effects of Nordic hamstring exercise on hamstring injuries in high school soccer players: a randomized controlled trial. Int J Sports Med 2020;41(3):154–160.

[28] Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, et al. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res Synth Methods 2019;10(1):83–98.

[29] IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp–Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014;14:25.

[30] Partlett C, Riley RD. Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation. Stat Med 2017;36(2):301–17.

[31] Wang CC, Lee WC. A simple method to estimate prediction intervals and predictive distributions: summarizing meta-analyses beyond means and confidence intervals. Res Synth Methods 2019;10(2):255–66.

[32] Viechtbauer W. Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med 2007;26(1):37–52.

[33] Sebelien C, Stiller CH, Maher S, Qu X. Effects of implementing Nordic hamstring exercises for semi-professional soccer players in Norway. Orthop Pract 2014;26:90–7.

[34] van der Horst N, Smits DW, Petersen J, Goedhart EA, Backx FJ. The preventive effect of the nordic hamstring exercise on hamstring injuries in amateur soccer players: a randomized controlled trial. Am J Sports Med 2015;43(6):1316–23.

[35] Petersen J, Thorborg K, Nielsen MB, Budtz-Jorgensen E, Holmich P. Preventive effect of eccentric training on acute hamstring injuries in men's soccer: a cluster-randomized controlled trial. Am J Sports Med 2011;39(11):2296–303.

[36] Gabbe BJ, Branson R, Bennell KL. A pilot randomised controlled trial of eccentric exercise to prevent hamstring injuries in community-level Australian Football. J Sci Med Sport 2006;9(1-2):103–9.

[37] Engebretsen AH, Myklebust G, Holme I, Engebretsen L, Bahr R. Prevention of injuries among male soccer players: a prospective, randomized intervention study targeting players with previous injuries or reduced function. Am J Sports Med 2008;36(6):1052–1060.

[38] Brooks JH, Fuller CW, Kemp SP, Reddin DB. Incidence, risk, and prevention of hamstring muscle injuries in professional rugby union. Am J Sports Med 2006;34(8):1297–306.

[39] Arnason A, Andersen TE, Holme I, Engebretsen L, Bahr R. Prevention of hamstring strains in elite soccer: an intervention study. Scand J Med Sci Sports 2008;18(1):40–8.

[40] Seagrave RA 3rd, Perez L, McQueeney S, Toby EB, Key V, Nelson JD. Preventive effects of eccentric training on acute hamstring muscle injury in professional baseball. Orthop J Sports Med 2014;2(6):2325967114535351.

[41] van der Horst N, Smits DW, Petersen J, Goedhart EA, Backx FJ. The preventive effect of the Nordic hamstring exercise on hamstring injuries in amateur soccer players: study protocol for a randomised controlled trial. Inj Prev 2014;20(4):e8.

[42] Elerian AE, El-Sayyad MM, Dorgham HMA. Effect of pre-training and post-training Nordic exercise on hamstring injury prevention, recurrence, and severity in soccer players. Ann Rehabil Med 2019;43(4):465–73.

[43] Lee H, Herbert RD, Lamb SE, Moseley AM, McAuley JH. Investigating causal mechanisms in randomised controlled trials. Trials 2019;20(1):524.

[44] Lee KJ, Thompson SG. Flexible parametric models for random-effects distributions. Stat Med 2008;27(3):418–34.

[45] Deeks JJ, Higgins JPT, Altman DG, Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. Chapter 10: Analysing data and undertaking meta-analyses. Cochrane handbook for systematic reviews of interventions, 6, Cochrane: version; 2021. updated February 2021.

[46] Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy 2002;7(1):51–61.

[47] Riley RD, Gates S, Neilson J, Alfirevic Z. Statistical methods can be improved within Cochrane pregnancy and childbirth reviews. J Clin Epidemiol 2011;64(6):608–18.

[48] Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1(2):97–111.

[49] Furlan AD, Pennick V, Bombardier C, van Tulder M. Editorial Board CBRG. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976) 2009;34(18):1929–41.

[50] Mack CD, Kent RW, Coughlin MJ, Shiue KY, Weiss LJ, Jastifer JR, et al. Incidence of Lower Extremity Injury in the National Football League: 2015 to 2018. Am J Sports Med 2020;48(9):2287–94.

[51] Whalan M, Lovell R, McCunn R, Sampson JA. The incidence and burden of time loss injury in Australian men's sub-elite football (soccer): A single season prospective cohort study. J Sci Med Sport 2019;22(1):42–7.

[52] Rees H, McCarthy Persson U, Delahunt E, Boreham C, Blake C. The burden of injury in field hockey: a secondary analysis of prospective cohort data. Scand J Med Sci Sports 2021;31(4):884–93.

[53] Roe M, Murphy JC, Gissane C, Blake C. Time to get our four priorities right: an 8-year prospective investigation of 1326 player-seasons to identify the frequency, nature, and burden of time-loss injuries in elite Gaelic football. PeerJ 2018;6:e4895.

[54] Owoeye OBA, VanderWey MJ, Pike I. Reducing injuries in soccer (football): an umbrella review of best evidence across the epidemiological framework for prevention. Sports Med Open 2020;6(1):46.