Implementing the 27 PRISMA 2020 Statement items for systematic reviews in the sport and exercise medicine, musculoskeletal rehabilitation and sports science fields: the PERSiST (implementing Prisma in Exercise, Rehabilitation, Sport medicine and Sports science) guidance

Clare L Ardern, Fionn Büttner, Renato Andrade, Adam Weir, Maureen C Ashe, Sinead Holden, Franco M Impellizzeri, Eamonn Delahunt, H Paul Dijkstra, Stephanie Mathieson, Michael Skovdal Rathleff, Guus Reurink, Catherine Sherrington, Emmanuel Stamatakis, Bill Vicenzino, Jackie L Whittaker, Alexis A Wright, Mike Clarke, David Moher, Matthew J Page, Karim M Khan, Marinus Winters

ABSTRACT

Poor reporting of medical and healthcare systematic reviews is a problem from which the sports and exercise medicine, musculoskeletal rehabilitation, and sports science fields are not immune. Transparent, accurate and comprehensive systematic review reporting helps researchers replicate methods, readers understand what was done and why, and clinicians and policy-makers implement results in practice. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement and its accompanying Explanation and Elaboration document provide general reporting examples for systematic reviews of healthcare interventions. However, implementation guidance for sport and exercise medicine, musculoskeletal rehabilitation, and sports science does not exist. The Pensa in Exercise, Rehabilitation, Sport medicine and Sports science (PERSiST) guidance attempts to address this problem. Nineteen content experts collaborated with three methods experts to identify examples of exemplary reporting in systematic reviews in sport and exercise medicine (including physical activity), musculoskeletal rehabilitation (including physiotherapy), and sports science, for each of the PRISMA 2020 Statement items. PERSiST aims to help: (1) systematic reviewers improve the transparency and reporting of systematic reviews and (2) journal editors and peer reviewers make informed decisions about systematic review reporting quality.

INTRODUCTION

High-quality systematic reviews deliver quality evidence to readers in a timely fashion—supporting informed decision making in practice.1 Poor reporting of medical and healthcare systematic reviews is prevalent, including in sport and exercise medicine, musculoskeletal rehabilitation and sports science.2,3 Systematic reviews must be clearly, transparently, accurately and comprehensively reported. There must be sufficient detail to allow researchers to replicate methods, for readers to understand what the systematic reviewers have done and why, and for clinicians/practitioners and policy-makers to act.

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Statement, published in March 2021, in five leading biomedical journals4–8—a substantially revised and updated version of the PRISMA 2009 Statement10—outlines the minimum items that should be reported in a systematic review to improve transparency and interpretation. The PRISMA Statement is recommended by the EQUATOR Network as the reporting guideline for systematic reviews. The PRISMA 2020 Statement4 and its accompanying Explanation and Elaboration13 document explain and elaborate general reporting examples for systematic reviews of healthcare interventions, but do not include implementation guidance specific for the context of sport and exercise medicine, musculoskeletal rehabilitation and sports science. Context-specific examples and elaboration are likely to enhance implementation.12 The PRISMA 2020 Statement7 focuses on systematic reviews of interventions, but it can be used as a basis for reporting systematic reviews of other types of research (eg, aetiology, prevalence, prognosis).10

SCOPE OF THE PRISMA IN EXERCISE, REHABILITATION, SPORT MEDICINE AND SPORTS SCIENCE GUIDANCE

Pensa in Exercise, Rehabilitation, Sport medicine and Sports science (PERSiST) aims to support systematic reviewers in the sport and exercise medicine (including physical activity), musculoskeletal rehabilitation (including physiotherapy and physical...
A systematic review employs specific, systematic methods of searching, selecting, assessing, collating and synthesising evidence to address a clearly formulated review question.44 Systematic review methods aim to minimise bias and maximise the practical relevance of the results to a broader spectrum of end users (eg clinicians, practitioners, patients, athletes, care givers, researchers and policy-makers).

Meta-analysis
Meta-analysis is a statistical technique for synthesising results when appropriate data (eg, effect estimates and their variances) are available, to yield a quantitative summary of the results.30

When making decisions about meta-analysis, consider the available data, methods and clinical differences (heterogeneity) among included studies.44 54 Data pooling using meta-analysis can provide a more precise estimate of treatment effect, diagnostic accuracy or prognosis because of greater statistical power when the results of multiple studies are combined.

Network meta-analysis
Traditional systematic reviews compare two different treatments (or compare one treatment to a control). Network meta-analysis is a statistical technique for synthesising and comparing results from more than two different interventions, even when the interventions have not been directly compared in a trial (the analysis builds a network of interventions based on interventions that have been directly and indirectly compared in different trials).31

Network meta-analysis is a more sophisticated analysis approach because it compares all available interventions, can produce a more precise effect estimate and can rank interventions from most to least effective.32 We recommend collaborating with a statistician to plan and conduct a network meta-analysis.

Individual participant data meta-analysis
Systematic reviewers typically pool aggregate data (eg, group means) extracted from included studies. Individual participant data (IPD) meta-analysis involves systematic reviewers collecting the original data for each participant included in the eligible studies, validating/checking and reanalysing the data.82

Compared with an aggregate data review, synthesising IPD data can substantially improve the quantity and quality of data, and ultimately produce a more robust synthesis of the field.82 We recommend collaborating with a statistician to plan and conduct an IPD meta-analysis.

How to use the PERSiST guidance
PERSiST complements the primary PRISMA 2020 Statement9 to help systematic reviewers in sport and exercise medicine, musculoskeletal rehabilitation, and sports science implement the PRISMA 2020 items in their research context. We recommend systematic reviewers use PERSiST alongside the PRISMA 2020 Statement4 and PRISMA 2020 Explanation and Elaboration11 when planning and reporting their systematic review; systematic reviewers may also find the METHODS MATTER statement13 helpful for general guidance on research design, methods and reporting.

PERSiST presents at least 1 exemplar illustrating exemplary reporting for each of the 27 PRISMA 2020 items8. Some examples were lightly edited for flow, including removing citations or web addresses. In some examples, we highlight additional considerations, share helpful resources and make suggestions for reporting systematic reviews in sport and exercise medicine, musculoskeletal rehabilitation, and sports science—these are boxes headed: Note for systematic reviewers.
Where relevant, we retained the exemplar’s citations or hyperlinks. Citations were renumbered to appear in the PERSiST guidance reference list. The examples are intended to guide systematic reviewers in sport and exercise medicine, musculoskeletal rehabilitation, and sports science regarding what to report—systematic reviewers will make decisions about how to present information (eg, whether to use tables, what information is appropriate for appendices/supplementary files) based on the review content and journal requirements.

**Terminology**

We adopted and followed the terminology of the PRISMA 2020 Explanation and Elaboration\(^1\) (box 1). Most systematic reviews use group level (aggregate) data (eg, mean and SD) from included studies. More sophisticated approaches to data analysis and synthesis include network meta-analysis, individual participant data (IPD) meta-analysis (box 1), umbrella reviews (sometimes called overviews or review of reviews) and prospective meta-analysis.\(^2\)

Systematic reviewers who are considering these more sophisticated approaches should visit the PRISMA website (www.prisma-statement.org) for guidance on reporting (including PRISMA Statement extensions for network meta-analysis,\(^3\) IPD meta-analysis,\(^4\) scoping reviews,\(^5\) diagnostic test accuracy,\(^6\) reporting of harms and health equity studies,\(^7\) and protocols.\(^8\) Other extensions are in development for: newborn and child health research (PRISMA), rapid reviews (PRISMA), ethically sensitive topics (PRISMA-Ethics), animal research (PRISMA Extension of Preclinical In Vivo Animal Experiments) and outcome measurement instruments (PRISMA-COSMIN).

**IMPLEMENTING THE PRISMA 2020 STATEMENT: EXAMPLES FOR SPORT AND EXERCISE MEDICINE, MUSCULOSKELETAL REHABILITATION, AND SPORTS SCIENCE**

**Title**

*Item 1: title*

Identify the report as a systematic review.

**Example:** ‘Comparative effectiveness of treatment options for plantar heel pain: a systematic review with network meta-analysis.’\(^9\)

**Example:** ‘Effectiveness of conservative interventions including exercise, manual therapy and medical management in adults with shoulder impingement: a systematic review and meta-analysis of randomised controlled trials (RCTs).’\(^10\)

**Abstract**

*Item 2*

See the PRISMA 2020 for Abstracts checklist (see list and examples in the summary box at the end of this document).

**Introduction**

*Item 3: rationale*

Describe the rationale for the review in the context of existing knowledge.

**Example:** ‘The International Olympic Committee, among others, has called for more diligence to safeguard the physiological development of the paediatric athlete. Performing a cardiac preparticipation evaluation within paediatric populations is controversial due to a lack of international consensus with regard to when, how and who should undertake such examinations. While data from the USA indicate that paediatric black athletes are particularly susceptible to sudden cardiac death (SCD), there is a general lack of understanding as to which factors (eg, physical growth, race and sex) have the potential to increase the likelihood of generating a false-positive diagnosis and unnecessary disqualification from competitive sport. Consequently, the distinction between paediatric athlete’s heart and cardiac pathology associated with SCD is especially important for this population.’\(^11\)

**Note for systematic reviewers.** Frame the objectives of systematic reviews of interventions according to the population, intervention, comparison, outcomes (PICO) framework. For a guide on framing other types of systematic reviews (eg, prognosis, diagnostic test accuracy) we recommend Munn et al.\(^12\)

**Methods**

*Item 5: eligibility criteria*

Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.

**Example:** ‘Types of studies’

Eligible RCTs were identified from systematic reviews investigating the effects of exercise therapy and published in the Cochrane database of Systematic Reviews. RCTs, cluster-randomised trials and randomised crossover studies were included if they compared an exercise therapy intervention with a non-exercising control treatment.

**Types of participants**

Studies that included participants with or without a medical condition were eligible, except for participants receiving chemotherapy, as all or nearly all these participants are anticipated to experience adverse events. Otherwise, no studies were excluded based on specific characteristics of the participants.

**Types of intervention**

Exercise therapy was the main intervention and each exercise session had to include active exercise therapy for at least 50% of the total time. Furthermore, the exercise could not be combined with any pharmacological, surgical or electrotherapeutic...
intervention. Besides strength/resistance, aerobic and neuromuscular exercise (defined as exercise interventions targeting sensorimotor deficiencies and functional stability), the following active exercise interventions were also included: dancing, running, cycling, QiGong and Tai Chi. However, interventions like whole body vibration, facial exercises, yoga, stretching or range of motion exercises or bladder training were excluded. There were no restrictions on the setting in which the exercise therapy was performed, that is, classes, gymnasium, etc.

Types of control intervention
Studies with comparators such as a non-exercise therapy control group, usual care, attention intervention, etc were included, as were nutraceuticals, placebo and education (eg, back school and similar interventions). However, studies where the control group involved any exercise (including stretching), pharmacological, surgical or electrotherapeutic intervention were excluded.

Type of outcome measures
The outcomes of interest were measures of adverse events. As classifying adverse events as treatment-related is largely subjective, and with unknown validity, the current study was not focused on reported adverse effects, but on adverse events as any undesirable event occurring during the study, divided into serious and non-serious adverse events.3

Methods
Item 6: information sources
Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.

Note for systematic reviewers. Ensure search decisions are tailored to the systematic review—avoid arbitrary decisions (eg, number of information sources to search).

For more information about grey literature, publication bias, and why it is important to search for grey literature in sport and exercise medicine, musculoskeletal rehabilitation, and sports science, see Winters and Weir31 and chapter 4 of the Cochrane Handbook (searching for and selecting studies).31 Searching grey (unpublished) literature identifies (1) ongoing studies that could be included when a systematic review is updated or that may warrant delaying publishing the review to ensure the study can be included, (2) completed studies that are unpublished and may never get published (eg, injury surveillance conducted by sports federations) and (3) books, monographs, dissertations, policy documents, reports etc that may address the systematic review question and provide relevant data. Reporting sources of grey literature will help readers determine the risk of bias due to missing (unpublished) studies.

Collaborate with a medical/healthcare librarian or information specialist—professionals who have extensive training in literature searching—when planning, developing and executing systematic review searches.32 Quality searching includes choosing appropriate databases and other information sources, designing search strategies for databases and registers, executing searches and saving and collating the results, documenting and reporting the search, and updating the search.33 For more information we recommend chapter 4 of the Cochrane Handbook.31

Example: ‘Electronic searches’
We [...] searched the following databases from inception to 11 March 2016 without restrictions to language or publication status:

► Cochrane Central Register of Controlled Trials (CENTRAL, which includes the Cochrane Back and Neck group (CBN) trials register) (The Cochrane Library, 2016, Issue 2).
► MEDLINE (OvidSP, 1946 to March week 1 2016; Appendix 1).
► MEDLINE In-Process & Other Non-Indexed Citations (OvidSP, 10 March 2016),
► Embase (OvidSP, 1980 to 2016 week 10),
► Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO, 1981 to 11 March 2016),
► PsycINFO (OvidSP, 2002 to March week 2 2016),
► Allied and Complementary Medicine Database (AMED) (OvidSP, 1985 to March 2016),
► CBN Trials Register (Cochrane Register of Studies (CRS)),
► Cochrane Complementary Medicine Field Trials Specialized Register (Cochrane Register of Studies Online (CRSO)),
► IndMED,
► PubMed,
► US National Institutes of Health ClinicalTrials.gov,
► World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP).

The searches were previously run in 2013 and 2014. In 2014, the ClinicalTrials.gov, WHO ICTRP and a supplementary search of the CBN Specialised Register in the CRS were added to the search strategy. In 2016, the PubMed search was revised to capture studies not in MEDLINE using the strategy recommended by Duffy 2014. The Information Specialist of the CBN conducted all searches except for the Cochrane Complementary Medicine Field Specialised Register, which we searched through the CRSO.

Searching other resources
We screened the reference lists of included studies and contacted experts in the field (eg, authors of included studies) for information on additional trials, including unpublished or ongoing studies.32

Example: ‘We conducted a comprehensive database search using PubMed, EMBASE, Cochrane Library, SPORTDiscus and PEDro to search clinical practice guidelines (CPGs) that presented rehabilitation of ACL injuries. To search grey literature and CPG repositories we used the OpenGrey, National Guideline Clearinghouse of the Agency for Healthcare Research and Quality, Guidelines International Network and National Institute for Health and Care Excellence (NICE) databases. [...] all searches up to the 30 September of 2018 [...]’.34

Methods
Item 7: search strategy
Present the full search strategies for all databases, registers and websites, including any filters and limits used.

Example: ‘Twelve systematic searches covering diagnosis, prevention and treatment for each of the four sections (1: hamstring, 2: adductor, 3: rectus femoris/quadriceps and 4: calf) [...]’. No restrictions were applied concerning year of publication, however, only publications in English were included. We
searched individual text words in title and abstract supplemented with Medical Subject Headings (MeSH) terms. We combined anatomical region of interest (eg, ‘Groin (MeSH)’ OR ‘adductor’ OR ‘groin’) AND type of injury (eg, ‘Athletic Injury (MeSH)’ OR ‘Strains and Sprains (MeSH)’ OR ‘strain’ OR ‘injur*’ OR ‘re-injur*’ OR ‘reinjur*’) AND outcome for diagnosis and treatment domains (eg, ‘Diagnosis (MeSH)’ OR ‘exam*’ AND ‘Return To Sport’ OR ‘full training,’ respectively) OR intervention for prevention domains (eg, ‘Primary Prevention (MeSHs)’ OR ‘Reduct*’). […] A flow chart of searches and the complete search strategy for all searches and databases is available as supplemental.38

METHODS

Item 8: selection process
Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.

Example: ‘The selection of studies was a three-stage process, with the identified citations independently evaluated for inclusion by two reviewers. The first stage was evaluation of titles selected with systematic searches described above. The article was included in this first screen if the title identified athletes and/or lumbar discectomy. We then reviewed the abstracts of all articles identified as meeting the search criteria. Full-text articles meeting criteria were retrieved and read independently by both reviewers and assessed for inclusion in the study. Disagreement was resolved by consensus between the two reviewers and a third reviewer if consensus could not be reached.’39

Table 1 Patient-important outcomes identified by the guideline panel aligned to the most corresponding outcome set available in the included trials and studies (reproduced with permission)

| Panel-identified outcomes | Outcomes chosen from included trials and studies |
|---------------------------|-----------------------------------------------|
| Pain                      | Pain (Visual Analogue Scale and Numeric Rating Scale, various scales) |
| Physical function         | Combined outcomes of physical function, physical capacity and pain items (Constant Score and its modifications, Shoulder Disability Questionnaire, Neer Score) |
| Global perceived effect   | Global perceived effect. The outcome was derived by subtracting the patients who reported ‘worse’ or ‘much worse’ from the no of patients reporting ‘much better’ or ‘no shoulder problems at all’/’healed completely’ at the relevant time points |
| Quality of life           | Health-related quality of life (EQ-5D-3L and 15D) |
| Participation             | Return to leisure activities/sport. Working status (working: yes/no) |
| Development of full-thickness rotator cuff tears | Prevalence of full-thickness rotator cuff tears at follow-up |
| Harms                     | Serious harms |

Note for systematic reviewers. The PRISMA Search Reporting Extension (PRISMA-S) and accompanying checklist35 aids quality reporting for search strategies. PRISMA-S outlines how and what to report so others can reproduce the search. Consider seeking peer review from an information specialist or librarian (eg, using the PRESS 2015 Guideline Evidence-Based Checklist36 when developing a search strategy to improve the quality of the search.37

Example: ‘Two reviewers independently extracted data using a specifically designed standardised data extracting form (see study protocol), and afterwards the reviewers compared the extracted data for consistency. All inconsistencies between the two forms were resolved by discussion between the two data extractors. Any disagreement between the data extractors after the initial discussion related to inconsistencies between the two individual data extractions was to be solved by involving a third person. General study information, participants and intervention characteristics, compliance, adverse events, withdrawals and outcome measures were extracted. Where data were not available from tables or the results section, the authors of the study in question were contacted by email, with one reminder after 2 weeks, if they did not respond to the first email.’40

Example: ‘Outcomes
The guideline panel identified outcomes of importance to patients; we used outcomes from the trials that most closely corresponded to those chosen by the patients (the systematic reviewers worked with 33 patient partners to identify and prioritise the outcomes of interest for the systematic review) (table 1). Because it was the most bothersome symptom for 86% of the patients who completed our survey, we considered pain as the most important outcome in this systematic review.

We defined serious harms as death, bleeding (uncontrolled or requiring transfusion), cardiac arrest requiring cardio-pulmonary resuscitation, myocardial infarction, cerebrovascular accident, acute renal failure, unplanned intubation, requiring ventilator for >48 hours, deep infection (surgical site or organ/space) sepsis, septic shock, pneumonia, wound dehiscence, pulmonary embolism, deep vein thrombosis or peripheral nerve injury. […] We used a priori-defined decision rules for data extraction: 1. When trialists reported final values and change from baseline values for the same outcome, we extracted final values. 2. When trialists reported unadjusted and adjusted values for the same outcome, we extracted the unadjusted values. 3. When trialists reported data based on the intention-to-treat (ITT) sample and another sample (eg, per-protocol, as-treated), we extracted ITT-analysed data.'
Data management
When trials used different outcome measures to evaluate the same construct, we chose the most common outcome measure as the index and transformed mean differences and SD of other outcome measures to the index instrument, and pooled the data using mean difference as the summary estimate. For trials not reporting the index instrument, we followed the data using mean difference as the summary estimate. When trials used different outcome measures to evaluate the same construct, we chose the most common outcome measure and transformed mean differences and SD of other outcome measures to the index and pooled as the index and transformed mean differences and SD of other outcome measures to the index instrument, and pooled the data using mean difference as the summary estimate. If multiple periods during which the pain was evaluated were available, the shortest was chosen.

Function outcomes and mixed function-capacity-pain scores
Oxford Shoulder Score.
Constant-Murley Score.
American Shoulder and Elbow Surgeons Standardised Form.
UCLA Shoulder Score.
Simple Shoulder Test.
Shoulder Disability Questionnaire.
Neer Score.
Watson-Sonnabend Score.
Health-related quality of life.
EQ-5D.
15D.
EQ-VAS.
SF-36.

Example of list of other variables for which data were sought:
‘We also extracted the following data: trial characteristics, patient demographic variables, diagnosis, treatment and data about trial methodology. Online supplementary appendix table 1 (box 3) presents a full list of extracted data items.’41

Methods
Item 10b: data items
List and define all other variables for which data were sought (eg, participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.

Note for systematic reviewers. If you have a long and/or detailed list of variables for which data were sought, consider including a summary of the data items (ie, variables other than outcome(s)) that were extracted as an appendix supplemental file or as a file uploaded to an online repository (eg, Open Science Framework).

Example of explanation for assumptions made about any missing or unclear information:
‘If authors did not report relevant numeric outcome data in the text, we contacted the authors or, when available, extracted the data from figures and graphs.’41
Note for systematic reviewers. For information on assessing risk of bias in the context of systematic reviews in the sport medicine, musculoskeletal rehabilitation, and sports science fields, we recommend Büttner et al. Some systematic reviewers use the terms ‘quality assessment’ and ‘risk of bias assessment’ interchangeably. These terms are not synonymous. Quality is poorly defined, but is often used to convey how well the research was conducted and reported. Bias refers to systematic deviations from the truth, which can occur due to flaws in research design, conduct, analysis and/or reporting (eg, outcome reporting biases, spin, per-protocol analyses).

We recommend systematic reviewers complete and report a risk of bias assessment using a clinimetric tool that is appropriate for the review question (eg, causation, prediction, diagnosis). Use a risk of bias assessment tool that facilitates domain-based assessment wherever possible. Do not calculate and present a numerical ‘methodological quality’ score when you mean to assess risk of bias (including quality assessment scales or composite reporting scales). When choosing a risk of bias assessment tool, carefully consider the ideal study design to answer the review question, and the key sources of bias that may influence the results of the systematic review.

The most relevant sources of bias differ depending on the study design. Assessing risk of bias requires a careful approach, tailored to the key threats to the internal validity (ie, bias) of the research question the systematic review aims to address. Different tools are appropriate for different study designs, and may include:

- The Cochrane ROB 2.0 for assessing bias in randomised controlled trials.
- ROBINS-I for assessing bias in non-randomised intervention studies.
- PROBAST for assessing bias in prediction modelling studies.
- QUIPS for assessing bias in prognostic studies.
- QUADAS-II for assessing bias in diagnostic accuracy studies (see PRISMA-DTA Statement).

Example: ‘When trials used different outcome measures to evaluate the same construct, we chose the most common outcome measure as the index and transformed mean differences and SDs of other outcome measures to the index instrument.’

The assessment was performed independently by two reviewers [...] The reviewers did not perform risk of bias assessment or data extraction for publications in which they were involved as an author. Disagreements were resolved via consensus or by a third reviewer [...] if necessary.45

Methods

Item 12: effect measures
Specify for each outcome the effect measure(s) (eg, risk ratio, mean difference) used in the synthesis or presentation of results.

Example: ‘Synthesis of results
For the analysis on benefits, we calculated the effect sizes in the individual studies as standardised mean differences, allowing pooling and comparison of the various outcomes assessed in the individual trials. We estimated the standardised mean difference as the difference between the mean score of the intervention and control groups divided by the pooled SD of the final score. This estimate of the effect size using standardised mean difference has a slight bias overestimating the effect size, and we applied a correction factor to convert the effect size to Hedges’ g.

In the analysis on harms, we transformed the numbers of adverse events into log odds of events, allowing pooling of data from the individual studies. Results are reported as number of adverse events per 1000 procedures with 95% CIs.46

Example: ‘We conducted meta-analyses, guided by considerations of bias. For the primary comparison, we pooled data from comparisons at low risk of bias. For the secondary comparison, we pooled data from comparisons irrespective of bias.

We assessed outcomes at 3 months, 6 months, 1 year, 2 years (for which we pooled data up to 3 years if no 2-year data were available), 5 years (we prioritised time points closest to 5 years) and >10 years following randomisation.

Example: ‘Studies were stratified by follow-up time categories: <2 years, 2 to 3 years, 5 to 10 years and >10 years [...]. Predefined subgroups were (1) patients treated with ACL reconstruction compared with non-operative treatment and (2) skeletally immature patients compared with skeletally mature patients. We accepted skeletal immaturity as defined in the study [...]. If skeletal immaturity was not defined in the study, we applied our definition of age under 16 years at injury for all patients.’

Methods

Item 13a: synthesis methods
Describe the processes used to decide which studies were eligible for each synthesis (eg, tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).

Example: ‘In studies with no ORs presented, the data were transformed to ORs from standard mean difference of muscle strength between the group of participants who developed osteoarthritis and those who did not. Data from adjusted analyses were extracted if available.’

Example: ‘When trials used different outcome measures to evaluate the same construct, we chose the most common outcome measure as the index and transformed mean differences and SDs of other outcome measures to the index instrument.’

Example: ‘When not reported, the SD was estimated from the SE of the mean, 95% CI, p value or other methods suggested in the Cochrane Handbook. Means and SDs were estimated from median and range for two studies. Following Cochrane guidelines, for any study that included two different intervention groups (ie, cycling vs walking) and one control group, the sample size in the control group was evenly divided so a comparison could be made to each intervention. We imputed r=0.5 when the correlation of prescores and postscores was required and we performed sensitivity analyses using r values ranging from 0.1 to 0.9.49

Methods

Item 13b: synthesis methods
Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics or data conversions.

Example: ‘In studies with no ORs presented, the data were transformed to ORs from standard mean difference of muscle strength between the group of participants who developed osteoarthritis and those who did not. Data from adjusted analyses were extracted if available.’

Example: ‘When trials used different outcome measures to evaluate the same construct, we chose the most common outcome measure as the index and transformed mean differences and SDs of other outcome measures to the index instrument.’

Example: ‘When not reported, the SD was estimated from the SE of the mean, 95% CI, p value or other methods suggested in the Cochrane Handbook. Means and SDs were estimated from median and range for two studies. Following Cochrane guidelines, for any study that included two different intervention groups (ie, cycling vs walking) and one control group, the sample size in the control group was evenly divided so a comparison could be made to each intervention. We imputed r=0.5 when the correlation of prescores and postscores was required and we performed sensitivity analyses using r values ranging from 0.1 to 0.9.49
following subgroups: woman, man and combined. A pooled estimate for the relative risk of ACL injury in women compared with men was calculated and summarised in a forest plot. Raw injury incidence rates for individual studies and pooled estimates were summarised in forest plots for the following groups and subgroups: woman, man, and combined. Pooled incidence rate ratios for women compared with men were calculated and summarised in forest plots.50

**Methods**

**Item 13d: synthesis methods**

Describe any methods used to synthesise results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.

**Example:** ‘Random effect models were used as large heterogeneity was expected due to the different approaches used to compare (ie, controls or contralateral leg) and assess knee extensor muscle strength (ie, isometric or isokinetic) as well as different pain and function scores. A standard [Cochran] Q-test was used to test the heterogeneity between studies, and the I² statistic measuring the proportion of variance attributable to study variance and calculated as the I² statistic measuring the proportion of variance attributable to inconsistency was subsequently calculated. […] I² each to 100% indicate maximal inconsistency between individual study results. Furthermore, the τ² value expressing the between study variance was estimated.51

**Example:** ‘Owing to expected clinical and methodological heterogeneity between the included studies, we used inverse variance random effects models to estimate relative risks with 95% CIs. […] We dealt with statistical heterogeneity using the I² statistic and prediction intervals. […] Analyses were conducted in either RevMan V.5.4 or Stata V.15.52

**Example:** ‘We performed meta-analyses using a random-effects model as heterogeneity was expected in participant, intervention and outcome characteristics. […] A random effects meta-analysis was applied to estimate the overall relative risk of adverse events in the exercise therapy groups compared with comparator groups. Heterogeneity was examined as between-study variance and calculated as the I² statistic measuring the proportion of variation in the combined estimates due to study variance. An I² value of 0% indicates no inconsistency between the results of individual trials, and an I² value of 100% indicates maximal inconsistency. Meta-analyses were performed in STATA (V.16.1) using the ‘meta’ command.53

**Note for systematic reviewers.** There are two common scenarios where systematic reviewers may wish to present a stratified analysis:

1. An analysis based on predefined groups when it is inappropriate for the specific review question to pool the data (eg, present injury incidence separately for women and men).
2. Discovering during the course of the review (eg, after extracting data) the need to stratify/subgroup data for a meaningful analysis (eg, short-term, medium-term and long-term outcomes of treatment for low back pain). This should not be confused with the expected and prespecified scenario for investigating statistical heterogeneity, where systematic reviewers conduct subgroup analyses to assess what is contributing to statistical heterogeneity.

**Methods**

**Item 13f: synthesis methods**

Describe any methods used to explore possible causes of heterogeneity among study results (eg, subgroup analysis, meta-regression).

**Example:** ‘We planned a sensitivity analysis for pain at 1 year to assess the impact of attrition bias due to missing data […]’. We also planned to assess small study bias by inspecting the distribution of funnel plots, but there were too few trials.41
Note for systematic reviewers. Use sensitivity analyses to check the robustness of decisions made when planning the data synthesis. Sensitivity analyses repeat a previous analysis to check whether a decision was reasonable (eg, the first analysis includes all studies; the sensitivity analysis includes only the studies at low risk of attrition bias). Wherever possible, prespecify sensitivity analyses in the systematic review protocol. Sometimes the need for sensitivity analyses only becomes apparent during the review process—ensure any unplanned sensitivity analyses are clearly labelled in the systematic review report.57

Example: ‘We controlled for smoking status and pre-existing diseases by performing additional sensitivity analyses. Analyses were restricted to never smokers, healthy participants, and healthy never smokers.’58

Example: ‘We examined the effects of methodological quality [bias] on the pooled estimate by removing studies that were at high or unclear risk of bias for the domains of blinding and incomplete outcome data. We had also intended to examine the effects of measurement device on the pooled estimate by removing studies that used pedometers, as previous studies suggest that these might be less accurate in detecting steps in people with COPD.’59

Example: ‘We planned to carry out sensitivity analyses on pain and function by examining the effects of: […]including trials with unclear allocation concealment (at risk of selection bias); including trials with an incomplete description of mild to moderate knee osteoarthritis; […] including trials at risk of detection bias (ie, unclear or no blinding of participant for participant-reported outcomes).60

Methods

Item 14: reporting bias assessment
Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).

Example: ‘In 15 of 16 included studies, the effect sizes were not reported; instead, the authors were contacted and asked to provide raw data. Therefore, the risk of publication bias was expected to be minor. Nevertheless, publication bias was assessed visually on a funnel plot (the effect size by the inverse of its standard error) and statistically using Egger’s test, the Begg and Mazumdar rank correlation test and the trim-and-fill method. It should be mentioned, however, that other factors, such as study quality and true heterogeneity, can produce asymmetry in funnel plots.’61

Methods

Item 15: certainty assessment
Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.

Example: ‘Data were synthesised and the quality of evidence were evaluated using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group. […] Two authors assessed the quality of evidence for each outcome relating to diagnostic tests (eg, effectiveness), prevention (eg, risk of injury) and treatment (eg, time to return-to-play) according to the approach from the GRADE working group. Agreement was reached by consensus. The quality of evidence was graded as: (1) high, indicating that further research is unlikely to change the confidence in the estimate of effect, (2) moderate, indicating that further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate, (3) low, indicating that further research is very likely have an important impact on the confidence in the estimate of effect and is likely to change the estimate or (4) very low, indicating high uncertainty about the estimate.

The starting quality of evidence was rated as ‘high’ when data were based on either RCTs for treatment and prevention purposes or rated as ‘low’ when based on observational studies. For diagnostic purposes, the starting quality of evidence was rated as high when based on cohort studies (prospective or cross-sectional). Subsequently, the quality of evidence could be downgraded one or two levels (eg, from high to moderate) for each of the following five domains of the GRADE approach: Study limitations (ie, serious risk of bias such as lack of blinding of outcome assessor or other concerns determined to influence the study result), inconsistency (ie, the heterogeneity of the results across studies if more than one study was included for the specific outcome), indirectness (ie, poor generalisability of the findings to the target population, eg, use groin injuries vs acute adductor injuries for prevention, and/or use of a clinically irrelevant outcome in relation to the question, eg, ‘time to end of treatment’ for ‘time to return-to-play’ outcomes), imprecision of the estimates (ie, wide CIs) and the risk of publication bias. Furthermore, the level of evidence for cohort studies could

Note for systematic reviewers. There are frameworks (eg, GRADE,62 CINeMA63) that help systematic reviewers provide outcome-based recommendations and judge the certainty (ie, how confident they are in the recommendation) of the body of evidence—supporting transparent critical appraisal and communication. We encourage using the descriptor ‘certainty’ instead of ‘quality’ when describing judgements.

Example: ‘Two authors assessed the quality of evidence for each outcome relating to diagnostic tests (eg, effectiveness), prevention (eg, risk of injury) and treatment (eg, time to return-to-play) according to the approach from the GRADE working group. Agreement was reached by consensus. The quality of evidence was graded as: (1) high, indicating that further research is unlikely to change the confidence in the estimate of effect, (2) moderate, indicating that further research is likely to have an important impact on confidence in the estimate of effect and may change the estimate, (3) low, indicating that further research is very likely to have an important impact on the confidence in the estimate of effect and is likely to change the estimate or (4) very low, indicating high uncertainty about the estimate.

The starting quality of evidence was rated as ‘high’ when data were based on either RCTs for treatment and prevention purposes or rated as ‘low’ when based on observational studies. For diagnostic purposes, the starting quality of evidence was rated as high when based on cohort studies (prospective or cross-sectional). Subsequently, the quality of evidence could be downgraded one or two levels (eg, from high to moderate) for each of the following five domains of the GRADE approach: Study limitations (ie, serious risk of bias such as lack of blinding of outcome assessor or other concerns determined to influence the study result), inconsistency (ie, the heterogeneity of the results across studies if more than one study was included for the specific outcome), indirectness (ie, poor generalisability of the findings to the target population, eg, use groin injuries vs acute adductor injuries for prevention, and/or use of a clinically irrelevant outcome in relation to the question, eg, ‘time to end of treatment’ for ‘time to return-to-play’ outcomes), imprecision of the estimates (ie, wide CIs) and the risk of publication bias. Furthermore, the level of evidence for cohort studies could
Consensus statement

Table 2 Criteria for downgrading the quality of outcomes using the grade approach (reproduced with permission)
therapy being delivered together with exercise or less than 50% of the intervention being exercise, 69 for the intervention involving vibration therapy, bladder training or other interventions not meeting the inclusion criteria for interventions. Further, 59 were excluded as they were reported in a language other than English, 45 because the full text was not accessible. Lastly, 8 were excluded because data were not extractable, 5 because the thesis was not available and 40 for other reasons.43

Results
Item 17: study characteristics
Cite each included study and present its characteristics.

Note for systematic reviewers. Report the key characteristics of the included studies to help readers understand how the included studies address the review question(s). The Template for Intervention Description and Replication (TIDieR) framework might help systematic reviewers summarise intervention details in intervention systematic reviews.

Example: ‘The final analysis included a total of 555 youth (283 and 272 in concurrent exercise and aerobic exercise training group, respectively). Three studies included the same population, but analysed different parameters. Nine studies recruited obese youth exclusively; whereas the rest targeted both overweight and obese children. Most studies (n=8) included adolescents (aged 13–18 years), one included only children, and one enrolled both children and adolescents. All studies included boys and girls. Sample sizes across studies ranged from 30 to 150, with a mean of 55 participants.

The primary mode of the aerobic exercise training programmes were based on treadmills and cycle ergometers, elliptical trainers, walking and running programmes, and sports participation. The exercise intensity was monitored using either maximum heart rate, or peak oxygen uptake.

For resistance exercise, studies used body weight exercise, free weights, selectorised machines or circuit training. Interventions duration varied from 10 to 48 weeks, with a mean of 30 weeks47 (figure 2 (labelled table 2 in the original publication) is an example of how one might present study characteristics).

Results
Item 18: risk of bias in studies
Present assessments of risk of bias for each included study.

Note for systematic reviewers. Please read in conjunction with item 11 explanation text.

Example: ‘Twenty-two trials (76%) were at high risk of bias. We had some concerns about bias in seven trials (24%). No trials were at low risk of bias. In studies that used both the VISA-A score and return to sport as outcome measures, there was no difference in risk of bias between the outcomes. In 48% of the trials, outcome measurement was a source of bias. All other sources of bias were also commonly judged as high risk: the randomisation procedure (21%), deviations from the intended intervention (28%), missing outcome data (28%) and selection of reported results (24%).43

Results
Item 19: results of individual studies
For all outcomes, present, for each study: (1) summary statistics for each group (where appropriate) and (2) an effect estimate and its precision (eg, CI/credible interval), ideally using structured tables or plots.

Example included in figure format: ‘One low quality study used x-ray to examine the pubic symphyses of athletes with hip/groin pain and those without pain. A reliable grading system quantified the abnormalities, which were present in all the hip/groin pain subjects (9/20 slight, 9/20 intermediate, 2/20 advanced). In contrast, the athletic control subjects had either no (3/20) or slight (17/20) abnormalities seen on x-ray. Another moderate quality study only reported the radiographic findings in the hip/groin pain subjects and not the control subjects. Synthesising the data on x-ray investigations of the pubic symphyses, there is currently limited evidence that x-ray findings differentiate athletes with hip/groin pain from athletes without pain.

Three moderate quality studies examined pubic bone oedema using MRI in athletes with hip/groin pain and controls. Dichotomous data for the presence or absence of bone oedema were extracted and pooled from these studies. The results indicated that there were high odds that participants with bone oedema on MRI would be in the hip/groin pain group with a large effect size; OR=41.63 (95% CI 1.6 to 1096.60). However, there was high heterogeneity demonstrated by this pooled result (I²=88%), and significant sensitivity to the data from the study of Cunningham et al (figure 3 (labelled figures 8 and 9 in the original publication)). The removal of this study data resulted in an OR=8.1 (95% CI 3.1 to 21.2) for the presence of bone oedema in subjects with hip/groin pain, representing moderate evidence, with a large effect size, that bone oedema in the pubic symphysis differentiates athletes with hip/groin pain from those without this pain.48

Results
Item 20a: results of syntheses
For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.

Note for systematic reviewers. For more information on outcome-based risk of bias assessment, please refer to Büttner et al.42 43 Read in conjunction with item 11 text. Perform separate risk of bias assessments for each outcome. Lumping all outcomes in a single summary risk of bias assessment (study-level assessment) risks spurious overall judgments because study limitations can distort different outcomes in different ways.

Example:
‘Risk of bias assessment
For the comparisons of subacromial decompression surgery versus placebo surgery, the risk of bias was low for all outcomes (figure 4). Due to detection bias (all studies), selection bias, attrition bias and selective reporting, for the comparisons of subacromial decompression surgery vs non-surgical treatment, the risk of bias was high for all outcomes except rotator cuff tears.

Primary comparisons
Two trials were at low risk of bias and sufficiently clinically and methodologically homogeneous to allow pooling for the comparison of subacromial decompression surgery plus postoperative rehabilitation versus placebo surgery plus postoperative rehabilitation.’41
### Consensus statement

#### Table 2 Characteristics of the studies included in the meta-analysis

| Study | Participants (n)/BMI (kg/m² or percentile) | Groups | Modality | Intervention duration (weeks) | Frequency (days/week) | Session (minutes) | Outcomes |
|-------|------------------------------------------|--------|----------|-------------------------------|-----------------------|------------------|----------|
| Ackel-D’Elia et al., 2014 | 48 (16 boys, 32 girls, 15–18 years of age)/≥95 | CE | Treadmill + free weights | 24 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| Alberg et al., 2015 | 150 (44 boys, 106 girls, 14–18 years of age)/≥p85 | CE | Treadmill, cycle ergometer or elliptical + free weights and machines | 22 | 2 | 20–40 | VO2max |
| Alberg et al., 2016 | 133 (25 boys, 108 girls, 14–18 years of age)/≥p85 | CE | Treadmill, cycle ergometer or elliptical + free weights and machines | 22 | 2 | 20–40 | VO2max |
| Sigal et al., 2014 | 42 (28 boys, 14 girls, 14–18 years of age)/≥95 | CE | Treadmill + free weights | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| Cambier et al., 2013 | 26 (10 boys, 16 girls, 14–18 years of age)/≥95 | CE | Treadmill + free weights and machines | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| Dámaso et al., 2014 | 116 (57 boys, 82 girls, 14–18 years of age)/≥95 | CE | Treadmill | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| de Mello et al., 2011 | 30 (20 boys, 10 girls, 14–18 years of age)/≥p95 | CE | Treadmill + free weights | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| de Piano et al., 2012 | 30 (20 boys, 10 girls, 14–18 years of age)/≥p95 | CE | Treadmill + free weights | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| Inoue et al., 2015 | 45 (17 boys, 28 girls, 14–18 years of age)/≥p95 | CE | Treadmill + free weights | 48 | 3 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, insulin, HOMA-IR, and leptin |
| Lee et al., 2010 | 36 (12–14 years of age)/≥p95 | CE | Circuit weight training | 10 | 2 | 60 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, and VO2max |
| Monteiro et al., 2015 | 32 (19 boys, 13 girls, 11–17 years of age)/sex and age BMI-10T | CE | Walking and running + own body weight | 20 | 3 | 50 | BM, BMI, BF (% and kg), LBM (% and kg), glucose, and VO2max |

*Applied different periodisation.

AE, aerobic training; BF, body fat; BM, body mass; BMI, body mass index; CE, concurrent exercise; CRP, C reactive protein; DBP, diastolic blood pressure; HDL, high-density lipoprotein; HOMA-IR, homeostatic model assessment; IOTF, International Obesity Task Force; LBM, lean body mass; LDL, low-density lipoprotein; SAT, subcutaneous fat; SBP, systolic blood pressure; TC, total cholesterol; TG, triglycerides; VAT, visceral fat; VO2max, maximal oxygen consumption; WC, waist circumference.

#### Figure 2
Example of study characteristics presented in a summary table (reproduced with permission).

#### Note for systematic reviewers
Present the study and population characteristics (eg, age, sex, n, etc.) for each synthesis separately. Below is an example of how to report study and population characteristics of contributing studies for each data synthesis.

**Example:**

Mental health symptoms and disorders among current elite athletes

Among those, 11 studies reported prevalence data on distress symptoms among 3335 male and female elite athletes (age ranging from 16 to 29 years) from team sports (eg, cricket, football, handball, ice hockey, rugby) and combined Olympic sports (eg, boxing, gymnastics, judo, rowing, swimming).

#### Results

**Item 20b: results of syntheses**

Present results of all statistical syntheses conducted. If meta-analysis was done, present each summary estimate and its precision (eg, CI/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.

**Example:**

‘Three studies reported on previous history of stress fracture and its association with increased risk of future stress...’
fracture. All three studies had similar findings in that athletes with a previous history of fracture were at increased risk of developing a future stress fracture with ORs ranging from 2.90 to 6.36. An exploratory meta-analysis confirmed the individual study results with runners with a previous history of stress fracture at five times higher risk of a future stress fracture (OR 4.99; 95% CI 2.91 to 8.56; p<0.001; I²=0%). Differences in risk between men and women did not reach statistical signif-

Example: ‘Stratified analyses for sex showed an increased risk in both men (OR 1.68, 95% CI 1.10, 2.58; I²=55.5%) and women (OR 1.59, 95% CI 0.94, 2.68; I²=54.3%). Differences in risk between men and women did not reach statistical significance (p=0.87).’

Results Item 20c: results of syntheses
Present results of all investigations of possible causes of heterogeneity among study results.

Example: ‘Publication bias
The funnel plot indicated that almost all studies fell within the expected parameters, most with low SE indicating that most studies were large. A majority of studies reported that women had greater incidence proportion than men. The funnel plot for incidence rate ratio indicated that most studies fell within the expected parameters. Standard error was relatively low, indicating that studies were large, and a majority of studies reported that women were at increased risk of ACL injuries relative to men. The studies are not evenly distrib-

Results Item 21: reporting biases
Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.

Example: ‘We assessed publication bias for this comparison. The funnel plot (figure 6) shows that small studies with negative effects are missing in the right lower quadrant, which indicates possible publication bias.’

Results Item 22: certainty of evidence
Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.
Table 3  Sensitivity analyses for the effect of multifactorial intervention on rate of falls and risk of sustaining one or more falls and recurrent falls

| Sensitivity analysis                              | Overall                                      |
|--------------------------------------------------|----------------------------------------------|
|                                                  | Rate of falls                                |
| Adequate concealment of allocation               | 10   4442                                    | 0.82 (0.68 to 0.97) | 94    |
| Adequate blinding of outcome assessors           | 15   8062                                    | 0.81 (0.71 to 0.92) | 92    |
| Adequate handling of incomplete outcome data*    | 12   5096                                    | 0.78 (0.68 to 0.89) | 90    |
| Individual randomisation                         | 18   6643                                    | 0.80 (0.70 to 0.91) | 91    |
|                                                  | Risk of sustaining one or more falls         |
| Adequate concealment of allocation               | 13   5296                                    | 0.95 (0.86 to 1.04) | 69    |
| Adequate blinding of outcome assessors           | 18   8481                                    | 0.95 (0.89 to 1.02) | 54    |
| Adequate handling of incomplete outcome data*    | 15   5732                                    | 0.96 (0.90 to 1.02) | 36    |
| Individual randomisation                         | 26   9839                                    | 0.95 (0.89 to 1.02) | 59    |
|                                                  | Risk of sustaining recurrent falls           |
| Adequate concealment of allocation               | 8    2589                                    | 0.89 (0.71 to 1.10) | 67    |
| Adequate blinding of outcome assessors           | 12   6668                                    | 0.89 (0.77 to 1.03) | 56    |
| Adequate handling of incomplete outcome data*    | 7    2129                                    | 0.96 (0.84 to 1.10) | 0     |
| Individual randomisation                         | 14   4095                                    | 0.89 (0.78 to 1.03) | 46    |

*Judgement of ‘low risk’ if there are no missing outcome data, or less than 20% of missing outcome data are missing and losses are balanced in numbers across intervention groups with similar reasons for missing data across groups or missing data have been imputed using appropriate methods.

Figure 5  Example of how to present results of sensitivity analyses (reproduced with permission).

Note for systematic reviewers. Critical appraisal frameworks (eg, GRADE, CINeMA) facilitate transparent critical appraisal and help systematic reviewers clearly communicate information. We encourage using the descriptor ‘certainty’ instead of ‘quality’ when describing judgements.

Example of GRADE summary reporting and outcome reporting in table format: ‘A summary of the quality of evidence, based on risk of bias, study design, CIs and variability in results, has been collated using the GRADE approach (Table 3).

All outcomes were rated as very low or low quality [we recommend systematic reviewers use the descriptor ‘certainty’ not ‘quality’] evidence demonstrating that the estimate of effect for those outcomes is uncertain.’

Discussion

Item 23a: discussion

Provide a general interpretation of the results in the context of other evidence.

Example: ‘There exist several other reviews, although previous reviews have focused on fewer interventions. The most important difference between our systematic review and previously published reviews is that we have a more stringent assessment of the risk of bias and quality of included trials. This is important because the strength of recommendations (eg, in future guidelines) will be based on the quality of the evidence.

For exercise, our results are in line with the other reviews, with the exception that we concluded that there is only very low-quality evidence where other studies reported moderate or even high or strong evidence. Two reviews evaluated scapula-focused treatments, reporting moderate evidence, and significant but clinically not relevant effects; whereas we did not separately analyse the scapula-focused treatments.’

Discussion

Item 23b: discussion

Discuss any limitations of the evidence included in the review.

Example: ‘Overall, there was a lack of consistent high-quality evidence to support nominating any particular movement quality outcome as a lower extremity injury risk factor due to inadequate reporting of concepts essential to establishing internal (how well an experiment was carried out) and external (can the results be applied to people and situations beyond the experiment) validity. The biggest threats to internal validity were related to the possible selection bias, and the reporting of, and adjustment for, potential influence of factors such as sex, injury history and training exposure.

Specifically, due to the lack of participant characteristic reporting, it was often difficult to determine if the individuals selected for a study differed systematically from those in
the source population (selection bias). Equally important was the consistent omission of the characteristics of those lost to follow-up, which made it impossible to determine if participants lost to follow-up were systematically different from those retained in a study. The inability to determine selection bias not only questions the internal validity of several included studies, it impacts the degree to which the findings of these studies can be generalised to the larger population from which the samples were drawn (external validity).\(^2\)

**Discussion**

**Item 23c: discussion**

Discuss any limitations of the review processes used; comment on the potential impact of each limitation.

**Example:** ‘We planned threshold analysis as a quantitative means to assess the robustness of network meta-analysis recommendations to potential limitations in the evidence. We were unable to use this approach because of substantial overlap in credible intervals from the network meta-analysis. Due to overlap in the intervals, no recommendations could be made, which is a fundamental prerequisite to performing a valid threshold analysis. To comply with our protocol, we report threshold results in (supplementary Web appendix), but chose to use GRADE to interpret the evidence. We were not able to evaluate small study bias due to too low number of trials. We found three completed trials in trial registers; two are under review, and the publication status of one trial is unknown.’\(^4\)

**Discussion**

**Item 23d: discussion**

Discuss implications of the results for practice, policy, and make recommendations for future research.

**Example:** ‘On average, patients with subacromial pain syndrome reported reduced pain, and improved physical function and quality of life following both surgical and non-surgical treatment. However, at up to 5 years, irrespective of treatment, patients continued to report pain of an average of 1.5–3 on a scale of 0–10 points on a Visual Analogue Scale. Clinicians working in primary care who are treating patients with subacromial pain syndrome should be aware that some patients experience prolonged symptoms and consider care strategies to support coping. A placebo control helps answer the research question “is there a benefit of subacromial decompression surgery?” because it minimises the risk of detection and performance biases. Both sources of bias contribute to overestimation of treatment effects by up to 20%. The largely consistent findings of the unblinded studies leave little doubt of the inference that subacromial decompression surgery provides no important benefit to patients. The current evidence provides no support for subacromial decompression surgery as an intervention providing important benefit for patients with subacromial pain syndrome. High-quality evidence indicates that surgery versus placebo surgery confers no important benefit on pain and function—the outcomes most important to patients. Considering the body of evidence, further head-to-head comparisons of subacromial decompression surgery compared with placebo surgery or non-surgical management with the same population are unlikely to change the results. Policymakers, funders and clinicians should consider these results in their funding and clinical decisions regarding the management of patients with shoulder pain.

Our review was designed to assess the benefits and harms of subacromial decompression surgery for managing subacromial pain syndrome. To date, no trial has demonstrated benefit of surgery for any clinical subgroup. In the future, subgroup claims should be supported by data from well-conducted trials at low risk of bias and the use of established criteria for credibility of subgroup effects, ideally enhanced by IPD meta-analysis.

For harms we welcome well-performed observational studies that specifically report harms after subacromial decompression surgery separately from harms following other types of shoulder surgery. Although the finding of no important benefit of surgery is robust, the root cause of subacromial pain and the underlying pathological process remain uncertain, as does the possible best treatment—if such exists—for subacromial pain syndrome. Network meta-analysis might provide information on this question and provide hypotheses to be tested in future methodologically sound trials. Future trialists investigating any treatment for subacromial pain syndrome should adopt a

---

**Table 3** Example of how to present assessments of certainty in the body of evidence for each outcome

| Outcome | Anticipated absolute effects (95% CI) | Risk with control group | Risk with resistance exercise training | No of participants (RCTs) | Certainty |
|---------|--------------------------------------|-------------------------|----------------------------------------|---------------------------|----------|
| ST      | 115.45 mm Hg                         | MD 3.17 mm Hg lower (6.95 lower to 0.60 higher) | 116 (4 RCTs) | ☢☢☢ VERY LOW *†‡ |
| ST      | 122.8 mm Hg                          | MD 4.02 mm Hg lower (5.92 lower to 2.11 lower) | 1456 (46 RCTs) | ☢☢☢ VERY LOW *†‡ |
| LT      | 131.6 mm Hg                          | MD 5.08 mm Hg lower (10.04 lower to 0.13 higher) | 366 (8 RCTs) | ☢☢☼ LOW*†‡ |
| ST      | 86.5 mm Hg                           | MD 3.31 mm Hg lower (6.86 lower to 0.78 higher) | 67 (3 RCTs) | ☢☢☢ VERY LOW*†‡ |
| MT      | 79.6 mm Hg                           | MD 1.57 mm Hg lower (4.60 lower to 1.46 higher) | 238 (10 RCTs) | ☢☢☢ VERY LOW*†‡ |
| MT      | 65.2 mm Hg                           | MD 0.72 mm Hg lower (3.66 lower to 2.22 higher) | 116 (4 RCTs) | ☢☢☢ VERY LOW *†‡ |
| MT      | 74.3 mm Hg                           | MD 1.73 mm Hg lower (2.88 lower to 0.57 lower) | 1418 (45 RCTs) | ☢☢☼ LOW*†‡ |
| LT      | 76.9 mm Hg                           | MD 4.93 mm Hg lower (8.58 lower to 1.28 lower) | 346 (7 RCTs) | ☢☢☢ VERY LOW*†‡ |

*Please refer to the original publication for the full results table; reproduced with permission.

†Downgraded due to potential for a recommendation or clinical course of action differing if the upper versus the lower boundary of the CI represented the truth and/or a sample size <400.

‡Publication bias suspected after inspection of funnel plots.

| Intermittent pain (NRS 0–10) | Risk with control group | Risk with resistance exercise training | No of participants (RCTs) | Certainty |
|------------------------------|-------------------------|----------------------------------------|---------------------------|----------|
| MT                           | 122.8 mm Hg             | MD 4.02 mm Hg lower (5.92 lower to 2.11 lower) | 1456 (46 RCTs) | ☢☢☢ VERY LOW *†‡ |
| LT                           | 131.6 mm Hg             | MD 5.08 mm Hg lower (10.04 lower to 0.13 higher) | 366 (8 RCTs) | ☢☢☼ LOW*†‡ |

*Inconsistent due to high heterogeneity, non-overlap of CI and/or markedly dissimilar point estimates.

†Publication bias suspected after inspection of funnel plots.
common set of outcomes, and the outcome measures should be standardised.\textsuperscript{44}
the majority of studies reported either 6-month or 12-month outcome data (not both). Therefore, to provide a more complete picture of early, medium and long-term outcomes, we modified the time point at which outcomes were report to: under 6 months, 6–12 months and over 12 months, respectively. One included study did not perform MRI in all patients prior to randomisation. 75

Example:

‘Deviations from study registration and study protocol

‘Agreement by raters on risk of bias decisions for the included randomised controlled studies was calculated as a percentage of agreement and κ values, and included in the results. Since the secondary analysis concerning type of programme showed that only the FIFA 11+ prevention programme was effective in reducing injuries, all secondary outcomes concerning lower limb, hamstring, knee and ankle injuries were only analysed in relation to this programme.

Furthermore, a post hoc analysis on hip/groin injury in relation to this programme was also included. Preplanned secondary analyses on the incidence rate ratio in the following subgroups: gender (male and female), and mean age groups (youth (<19 years), seniors (19–30 years), old girls/boys (31–39 years) and veterans (>39 years)) were not conducted, as the included studies did not allow making meaningful comparisons with only six studies, where studies with male (n=3) and female participants (n=3) significantly differed in the age group they targeted. The predefined secondary analysis of compliance at team level was not performed as all team-level data could not be obtained from the corresponding authors of the included studies. Instead, the preplanned analysis of the association between prevention programme compliance and injury incidence was further supported by a post hoc analysis of the association between prevention programme compliance and the overall injury incidence rate ratio from each study to accommodate for the risk of substantial variance in injury incidence between studies due to other factors than the FIFA injury prevention programmes. 40

Other information

Item 25: support
Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.

Example: ‘This research received no grant from any funding agency in the public, commercial or not-for-profit sectors. […] is supported and funded by the National Osteoporosis Society via the Linda Edwards Memorial PhD Studentship.’76

Item 26: competing interests
 Declare any competing interests of review authors.

Example: ‘All authors have completed the ICMJE uniform disclosure form at (available on request from the corresponding author) and declare: […] has received personal fees from Össur, Flexion Therapeutics, Medivir, Teijin, MerckSerono, Allergan, and Galapagos and is editor-in-chief of Osteoarthritis and Cartilage; […] has received personal fees for lectures and royalties for books from Össur, Finnish Orthopedic Society, Studentlitteratur, and Munksgaard and is an associate editor of Osteoarthritis and Cartilage; no other relationships or activities that may appear to have influenced the submitted work.’46

PERSPECTIVES

All PRISMA 2020 items apply to systematic reviews in the sport and exercise medicine, musculoskeletal rehabilitation, and sports science fields. However, some items require more elaboration to promote implementation. PERSiST is intended to support systematic reviewers to implement PRISMA 2020 in their systematic reviews in the sport and exercise medicine, musculoskeletal rehabilitation, and sports science fields, and to promote transparent reporting. Use PERSiST in parallel with the PRISMA 2020 Checklist4 and Explanation and Elaboration, 11 and with the METHODS MATTER statement17 (as appropriate). We encourage journal editors and reviewers in the relevant fields to use PERSiST to help them make informed judgements about the quality and transparency of systematic review reporting.

Researchers make many decisions at the various stages of conducting a systematic review—from planning the review, completing the search and analysis, and writing the report.78 Decisions include choosing search terms, which data to extract, data pooling policies, choice of statistical tests, among others.79 Poor reporting makes it impossible for other researchers to replicate a systematic review, and a systematic review that cannot be replicated has little value. PERSiST and the PRISMA 2020 Statement and Checklist4 will help systematic reviewers deliver quality systematic reviews in sport and exercise medicine, musculoskeletal rehabilitation, and sports science.
Consensus statement

Summary box. Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 for Abstracts checklist and accompanying examples

Item 1: Identify the report as a systematic review.
Example: ‘Effect of soft braces on pain and physical function in patients with knee osteoarthritis: systematic review with meta-analyses.’

Item 2: Provide an explicit statement of the main objective(s) or question(s) the review addresses.
Example: ‘To systematically review and synthesize the effect of soft braces on pain, and self-reported and performance-based physical function in patients with knee osteoarthritis.’

Item 3: Specify the characteristics used as criteria for eligibility.
Example: ‘Randomised controlled trials examining the effectiveness of any treatment in patients with both insertional and/or midportion Achilles tendinopathy. We excluded trials with N ≤10 per treatment arm or investigating tendon ruptures.’

Item 4: Specify the information sources (eg, databases, registers) used to identify studies and the date when each was last searched.
Example: ‘The following electronic databases were searched from inception to April 20, 2016: The Cochrane Central Registry for Controlled Trials (CENTRAL), PubMed, EMBASE, Cumulative Index to Nursing and Allied Health Literature, SPORTDiscus, Web of Science and PEDro.’

Item 5: Specify the methods used to assess risk of bias in the included studies.
Example: ‘Two reviewers independently extracted data and assessed risk of bias with Risk of Bias Tool V2. We used Grading of Recommendations, Assessment, Development and Evaluation to appraise the strength of the evidence.’

Item 6: Specify the methods used to present and synthesise results.
Example: ‘We pooled data for meta-analyses by length of follow-up, reported as mean differences or standardised mean differences using random-effects wherever possible, or the fixed-effect model, where appropriate. If a meta-analysis was not possible, we synthesised studies narratively.’

Item 7: Give the total number of included studies and participants, and summarise relevant characteristics of studies.
Example: ‘Twenty studies comprising 2375 injuries from 1234 athletes (all males and mean age of 24 years) from different sports were included. Internal (65%) and external loads (70%) were collected in more than half of the studies and the session-rating of perceived exertion and total distance were the most commonly collected metrics. The acute chronic workload ratio was commonly calculated using the coupled method (95%), 1:4 weekly blocks (95%) and subsequent week injury lag (80%). There were 14 different binning methods with almost none of the studies using the same binning categories.’

Item 8: Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (ie, which group is favoured).
Example: ‘Overall, 42.7% (95% confidence interval (CI): 18%, 69%) of patients passed return to sport (RTS) criteria, and 14.4% (95% CI: 8%, 21%) of those who passed experienced a second ACL injury (graft rupture or contralateral ACL injury). There was a nonsignificant 3% reduced risk of a second ACL injury after passing RTS criteria (risk difference, −3%; 95% CI: −16%, 10%; I²=74%, p=0.61).’

Item 9: Provide a brief summary of the limitations of the evidence included in the review (eg, study risk of bias, inconsistency and precision).
Example: ‘The evidence rating of the Grading of Recommendations Assessment, Development and Evaluation scale was ‘very low certainty’ due to imprecision and heterogeneity of the pooled risk difference estimate.’

Item 10: Provide a general interpretation of the results and important implications.
Example: ‘In our living network meta-analysis no trials were at low risk of bias and there was large uncertainty in the comparative estimates. For midportion Achilles tendinopathy, wait-and-see is not recommended as all active treatments seemed superior at 3-month follow-up. There seems to be no clinically relevant difference in effectiveness between different active treatments at either 3-month or 12-month follow-up. As exercise therapy is easy to prescribe, can be of low cost and has few harms, clinicians could consider starting treatment with a calf-muscle exercise programme.’

Item 11: Specify the primary source of funding for the review.
Example: ‘Funding: This research received a grant from the Dutch Association of Medical Specialists to develop a clinical guideline for the treatment of patients with Achilles tendinopathy. The Dutch Patient Federation is involved in this guideline development and assisted in sending out patient surveys.’

Competing interests: […] led a research project in collaboration with Pfizer (project ended 31 December 2018). Pfizer part-funded a junior researcher. The projects were purely methodological, using historical data on pharmacological treatments for pain relief.

Example: ‘Funding: The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.’

Item 12: Provide the register name and registration number.
Example: ‘Prospero registration: CRD42018086467.’

Item 13: Provide a statement of funding and competing interests.
Example: ‘Funding: This research received a grant from the Dutch Association of Medical Specialists to develop a clinical guideline for the treatment of patients with Achilles tendinopathy. The Dutch Patient Federation is involved in this guideline development and assisted in sending out patient surveys.’

Competing interests: […] led a research project in collaboration with Pfizer (project ended 31 December 2018). Pfizer part-funded a junior researcher. The projects were purely methodological, using historical data on pharmacological treatments for pain relief.

Example: ‘Funding: The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.’

Item 14: Provide a statement of any additional support or assistance provided for the review.
Example: ‘Prospero registration: CRD42018086467.’
Author affiliations
1 Musculoskeletal & Sports Injury Epidemiology Center, Sophiahemmet University, Stockholm, Sweden
2 Sport and Exercise Medicine Research Centre, La Trobe University, Melbourne, Victoria, Australia
3 Department of Family Practice, University of British Columbia, Vancouver, British Columbia, Canada
4 School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland
5 Clinical Research, Clínica do Dragão, Eschipregueira-Mendes Sports Centre - FIFA Medical Centre of Excellence, Porto, Portugal
6 Dom Henrique Research Centre, Porto, Portugal
7 Aspetar Sports Groin Pain Centre, Aspetar Orthopaedic & Sports Medicine Hospital, Doha, Qatar
8 Department of Orthopaedics and Sports Medicine, Erasmus MC University Medical Center, Rotterdam, The Netherlands
9 Sport Medicine and Exercise Clinic Haarlem (58K), Haarlem, The Netherlands
10 Centre for Hip Health and Mobility, University of British Columbia, Vancouver, British Columbia, Canada
11 SM, Department of Health Science and Technology, Aalborg University, Aalborg, Denmark
12 Center for General Practice, Aalborg University, Aalborg, Denmark
13 Faculty of Health, University of Technology Sydney, Sydney, New South Wales, Australia
14 Sports Medicine, Aspetar Orthopaedic and Sports Medicine Hospital, Doha, Qatar
15 Department for Continuing Education, University of Oxford, Oxford, UK
16 Institute for Musculoskeletal Health, Sydney School of Public Health, The University of Sydney, Sydney, New South Wales, Australia
17 Academic Center for Evidence Based Sports Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands
18 Charles Perkins Centre, School of Health Sciences, University of Sydney, Sydney, New South Wales, Australia
19 University of Queensland School of Health and Rehabilitation Sciences: Physiotherapy, The University of Queensland, Brisbane, Queensland, Australia
20 Department of Physical Therapy, Faculty of Medicine, The University of British Columbia, Vancouver, British Columbia, Canada
21 Arthritis Research Canada, Vancouver, British Columbia, Canada
22 Doctor of Physical Therapy Program, Tufts University School of Medicine, Tufts University, Boston, Massachusetts, USA
23 School of Medicine, Dentistry and Biomedical Sciences, Centre for Public Health, Queen’s University, Belfast, UK
24 Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
25 School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia
26 School of Kinesiology, University of British Columbia, Vancouver, British Columbia, Canada
27 Canadian Institutes of Health Research-Institute of Musculoskeletal Health, Vancouver, British Columbia, Canada
28 School of Kinesiology, University of British Columbia, Vancouver, British Columbia, Canada
29 Canadian Institutes of Health Research-Institute of Musculoskeletal Health, Vancouver, British Columbia, Canada

Twitter Claire L Ardem @claire_ardem, Fionn Büttrn @peanutbutterb, Renato Andrade @Renato_Physo, Sinedin Holden @Sinedin_Holden, Franco M Impellizzeri @francompwell, Eamonn Delahunth @EamonnDelahunth, H Paul Dijkstra @DPaulDijkstra, Stephanie Mathiessn @DSMathiessn, Catherine Sherrington @CathieS, Emmanuel Stamatakis @EM_Stamatakis, Bill Vicenzino @Bill_Vicenzino, Jackie L Whittaker @jwhittak_physio, Karim M Khan @KarimKhan_IMHA and Marinus Winters @marinuswinters

Contributors CLA, MW, FB and RA proposed the idea for PERSIST, and planned and coordinated the project. The PERSIST Working Group is: RA, MCA, FCA, FB, ED, HPD, SH, FMI, KMK, SM, MSR, GR, CS, ES, BV, AW, JLW, MW and AAW. RA, MCA, SH, FMI and AW led and coordinated the PERSIST Working Group teams in identifying and appraising examples. The PERSIST advisory panel is: MC, DM and MIP. CLA and MW wrote the first draft of the manuscript; all authors contributed to reviewing, editing and revising the manuscript, and approved the final submitted version. CLA is the project lead and guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests RA, MCA, FCA, FB, MC, SH, FMI, SM, MSR, GS, BV, JLW, MW and AAW declare they have no competing interests. CLA was a Deputy Editor (Systematic Reviews) for BJSM from 2016 to 2018. ED, HPD and AW are associate editors for BJSM. MCA was a member of the BJSM editorial board from 2008 to 2020. KMK was Editor-in-Chief of BJSM from 2008 to 2020. He holds no position with the BJSM or the BMJ Group at present (September 2021). DM is Chair of the PRISMA group, led the PRISMA 2009 statement and co-led the PRISMA 2020 statement. MJP co-led the PRISMA 2020 statement. ES was editor of BJSM from 2017 to 2020, and editor-in-chief of BJM Open Sport & Exercise Medicine from 2019 to 2020. He is a senior adviser to BJSM (September 2021).

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID IDs Claire L Ardem http://orcid.org/0000-0001-8102-3631
Fionn Büttrn http://orcid.org/0000-0002-5987-3063
Renato Andrade http://orcid.org/0000-0002-7636-7816
Adam Weir http://orcid.org/0000-0003-0861-662X
Maureen C Ashe http://orcid.org/0000-0002-6820-4425
Sinead Holden http://orcid.org/0000-0002-7314-2152
Franco M Impellizzeri http://orcid.org/0000-0002-1703-2573
Eamonn Delahunth http://orcid.org/0000-0001-5449-5932
H Paul Dijkstra http://orcid.org/0000-0003-3166-1357
Stephanie Mathiessn http://orcid.org/0000-0002-7335-8842
Catherine Sherrington http://orcid.org/0000-0001-8934-4368
Emmanuel Stamatakis http://orcid.org/0000-0001-7323-3225
Bill Vicenzino http://orcid.org/0000-0003-0253-5933
Jackie L Whittaker http://orcid.org/0000-0002-6591-4976
Alexis A Wright http://orcid.org/0000-0002-8429-0775
Mike Clarke http://orcid.org/0000-0002-2926-7257
Karim M Khan http://orcid.org/0000-0002-9976-0258
Marinus Winters http://orcid.org/0000-0001-5742-7441

REFERENCES
1 Sadett De, Rosenberg WM, Gray JA, et al. Evidence based medicine: what is it and what is it not? BMJ 1996;312:71–2.
2 Pussegoda K, Turner L, Gagnit C, et al. Systematic review adherence to methodological or reporting quality. Syst Rev 2017;6:131.
3 Page MJ, Shamseer L, Altman DG, et al. Systematic review characteristics of systematic reviews of biomedical research: a cross-sectional study. PLoS Med 2016;13:e1002028.
4 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
5 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 2021;88:105906.
6 Page MJ, McKenzie JE, Bossuyt PM. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. J Clin Epidemiol 2021;134.
7 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021;10:89.
8 Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. PLoS Med 2021;18:e1003583.
9 Page MJ, McKenzie JE, Bossuyt PM, et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol 2021;143:103–12.
10 Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.
11 Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021;372:n160.
12 Blom JW, Muth C, Glazioso F, et al. Describing deprescribing trials better: an elaboration of the CONSORT statement. J Clin Epidemiol 2020;127:87–95.
13 Nielsen Rasmus Østergaard, Shrier I, Casals M, et al. Statement on methods in sport injury research from the first methods matter meeting, Copenhagen, 2019. J Orthop Sports Phys Ther 2020;50:226–33.
14 Ioannidis J. Next-generation systematic reviews: prospective meta-analysis, individual-level data, networks and umbrella reviews. Br J Sports Med 2017;51:1456–8.
15 Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanation. Ann Intern Med 2015;162;777–84.
16 Stewart LA, Clarke M, Rovers M, et al. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. JAMA 2015;313:1657–65.
17 Trinco AC, Lillie E, Zarif W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169:467–73.
18 McInnes MDF, Moher D, Thoms BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 2018;319:388–96.
Consensus statement

19 Zorzaela L, Lake YK, Ioannidis JP, et al. PRISMA harms checklist: improving harms reporting in systematic reviews. BMJ 2016;352:i157.

20 Welch V, Petticrew M, Tugwell P, et al. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. PLoS Med 2012;9:e1001333.

21 Welch V, Petticrew M, Petrovic J, et al. Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E) 2012: explanation and elaboration. Int J Equity Health 2015;14:92.

22 Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1.

23 Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;350:g6742.

24 Bababude OQ, Legha A, Littlewood C, et al. Comparative effectiveness of treatment options for plantar heel pain: a systematic review with network meta-analysis. Br J Sports Med 2019;53:182–94.

25 Steuri R, Sattelmayer M, Eligia S, et al. Effectiveness of conservative interventions including exercise, manual therapy and management in adults with shoulder impingement: a systematic review and meta-analysis of RCTs. Br J Sports Med 2017;51:1340–7.

26 McClean G, Riding NR, Ardern CL, et al. Electrical and structural adaptations of the paediatric athlete’s heart: a systematic review with meta-analysis. Br J Sports Med 2018;52:230.

27 Munn Z, Stern C, Aromataris E, et al. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMJ Med Res Methodol 2018;18:S.

28 Karanasios S, Karkos J, Whiteley R, et al. Exercise interventions in lateral elbow tendinopathy have better outcomes than passive interventions, but the effects are small: a systematic review and meta-analysis of 2122 subjects in 30 trials. Br J Sports Med 2021;55:477–85.

29 Niesmeier A, Lund H, Stafne SN, et al. Adverse events of exercise therapy in randomised controlled trials: a systematic review and meta-analysis. Br J Sports Med 2020;54:1073–80.

30 Winters M, Weir A. Grey matters; on the importance of publication bias in systematic reviews. Br J Sports Med 2017;51:1488–9.

31 Leufvén C, Glanville J, Briscoe S. Chapter 4: Searching for and selecting studies. In: Lefebvre C, Glanville J, Briscoe S. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, eds. Cochrane Handbook for Systematic Reviews of Interventions version 6.1 (updated September 2020). Cochrane, 2020. www.training.cochrane.org/handbook.

32 Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, et al. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. J Clin Epidemiol 2015;68:617–26.

33 Wieland LS, Skoetz N, Pilkington K, et al. ‘What’s my risk of sustaining an ACL injury while playing football (soccer)?’ a systematic review with meta-analysis. Br J Sports Med 2017;51:702–7.

34 Nejstgaard CH, Bero L, Hrobjartsson A, et al. Association between conflicts of interest and favourable recommendations in clinical guidelines, Advisory Committee reports, opinion pieces, and narrative reviews: systematic review. BMJ 2020;371:m234.

35 Bricca A, Harris LK, Langer M, et al. Benefits and harms of exercise therapy in people with multimorbidity: a systematic review and meta-analysis of randomised controlled trials. Ageing Res Rev 2020;63:101166.

36 Grindem H, Mansournia MA, Øiestad BE, et al. Was it a good idea to combine the studies? Why clinicians should care about heterogeneity when making decisions based on systematic reviews. Br J Sports Med 2019;53:399–401.

37 Rethlefsen ML, Øiestad BE, Hart HF, et al. Prevalence of knee osteoarthritis features on magnetic resonance imaging in asymptomatic uninjured adults: a systematic review and meta-analysis. Br J Sports Med 2019;53:1268–78.

38 Hopewell S, Cospay B, Nicolaou P, et al. Multifactorial interventions for preventing falls in older people living in the community: a systematic review and meta-analysis of 41 trials and almost 20 000 participants. Br J Sports Med 2020;54:1340–50.

39 Deeks J, Higgins JP, Altman DG. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for systematic reviews of interventions version 6.1 (updated September 2020). Cochrane, 2020. www.training.cochrane.org/handbook.

40 Jayed A, Soltani S, Zargar MS, et al. Central fatness and risk of all cause mortality: systematic review and dose-response meta-analysis of 72 prospective cohort studies. BMJ 2020;370:m334.

41 Ruge AT, Cox NS, Abrams MJ, et al. Interventions to promote physical activity in people with chronic obstructive pulmonary disease (COPD). Cochrane Database Syst Rev 2020;4:CD012626.

42 Palmer JS, Monk AP, Hopewell S, et al. Surgical interventions for symptomatic mild to moderate knee osteoarthritis. Cochrane Database Syst Rev 2019;7:CD012128.

43 Papasavvas T, Bonov RO, Alhashemi M, et al. Depression symptom severity and cardiopulmonary fitness in healthy and depressed adults: a systematic review and meta-analysis. Sports Med 2016;46:219–30.

44 In: Schieven G, Brozek J, Guyatt G, et al, eds. GRADE handbook for grading quality of evidence and strength of recommendations. The GRADE Working Group, 2013. https://dgt.graderpo.org/app/handbook/handbook.html.

45 Nikolkopoulos A, Higgins JP, Papakonstantinou T, et al. CINeMA: an approach for assessing confidence in the results of a network meta-analysis. PLoS Med 2020;17:e1003082.

46 Ashton RE, Twe G, Aning JL, et al. Effects of short-term, medium-term and long-term resistance exercise training on cardiometabolic health outcomes in adults: systematic review with meta-analysis. Br J Sports Med 2020;54:341–8.

47 Grassi A, Ardern CL, Marchegiani Mucci GM, et al. Does revision ACL reconstruction measure up to primary surgery? A meta-analysis comparing patient-reported and clinician-reported outcomes, and radiographic results. Br J Sports Med 2016;50:716–24.

48 Hoffmann TC, Glasziou PP, Bounton J, et al. Better reporting of interventions: template for intervention description and replication (TIDier) checklist and guide. BMJ 2014;348:g1867.

49 Garcia-Hermoso A, Ramirez-Velez R, Ramirez-Campillo R, et al. Concurrent aerobic plus resistance exercise versus aerobic exercise alone to improve health outcomes in paediatric obesity: a systematic review and meta-analysis. Br J Sports Med 2018;52:161–6.
