Sports Medicine
https://doi.org/10.1007/s40279-025-02187-5

SYSTEMATIC REVIEW

# Physical Therapies for Delayed-Onset Muscle Soreness: An Umbrella and Mapping Systematic Review with Meta-meta-analysis

**Szczepan Wiecha[1] · Igor Cieśliński[2] · Paweł Wiśniowski[2] · Maciej Cieśliński[2] · Wojciech Pawliczek[2] · Paweł Posadzki[3,4] · Robert Prill[5,6] · Joanna Zając[7] · Maciej Płaszewski[2]**

Accepted: 4 February 2025


## Abstract

**Background** Delayed-onset muscle soreness commonly arises from intense and unaccustomed physical exercise, leading to reduced muscle strength, increased pain and inflammation. A number of systematic reviews evaluating physiotherapeutic treatments for delayed-onset muscle soreness have been published since the 1990s. However, these systematic reviews frequently yield conflicting findings, further impeding clinical practice.

**Objectives** The primary aim of this study was to summarise the effectiveness of physiotherapy interventions in alleviating delayed-onset muscle soreness through an umbrella review. Additionally, we evaluated the risk of bias in systematic reviews, synthesised their findings, and categorised the evidence strength to provide practical insights for clinicians and researchers.

**Methods** An umbrella review with a meta-meta-analysis was conducted. MEDLINE, Embase, Cochrane Database of Systematic Reviews, CINAHL, PEDro and Epistemonikos were searched from 1998 to February 2024. Systematic reviews of randomised controlled trials of any treatments used post-exercise by physiotherapists to reduce delayed-onset muscle soreness in healthy adults, regardless of their physical activity, were eligible. A MeaSurement Tool to Assess systematic Reviews-2 (AMSTAR-2) was used to evaluate the methodological quality of the included systematic reviews. Corrected covered areas were calculated to address the overlap of primary trials in the included systematic reviews. An evidence map was created to categorise and visualise the effects of interventions using a multi-dimensional approach, based on the effect size and strength of evidence (Class I–V), i.e. the number of cases, Hedges' *g*, *p*-value, heterogeneity, Egger's test and excess of significance bias test.

**Results** Twenty-nine systematic reviews with 863 unique randomised controlled trials, addressing 24 distinct physiotherapeutic treatments, met the inclusion criteria. Seventeen systematic reviews were of critically low methodological quality, with only two rated as high quality. The evidence map suggests significant effects in pain reduction immediately post-exercise for contrast therapy (Class II), cooling therapy and cryostimulation (Class IV); 24 h: massage therapy (Class III) and cooling therapy, contrast therapy, electrical stimulation, cryostimulation, phototherapy, heat therapy (Class IV); 48 h: compression, contrast therapy, kinesiotaping and cryostimulation (Class III) and cooling therapy, massage, phototherapy, heat therapy (Class IV); 72 h: kinesiotaping (Class III) and contrast therapy, cooling therapy, massage, phototherapy, vibration (Class IV); 96 h: compression, phototherapy, and contrast therapy (Class IV). The effect sizes (Hedges' *g*) ranged from 0.36 (95% confidence interval 0.46, 3.18) for cooling therapy to 1.82 (95% confidence interval 0.46, 3.18) for heat therapy indicating small and large effects, respectively.

**Conclusions** There is a large body of evidence from predominantly low-quality systematic reviews of randomised controlled trials evaluating the effectiveness of physiotherapeutic treatments for delayed-onset muscle soreness. There is some strong evidence to support the effectiveness of cooling therapy, cryostimulation, contrast therapy, massage, phototherapy and kinesiotaping at various follow-up intervals, whereas evidence for stretching, exercises and electrical stimulation is weak. Uncertainties, heterogeneity and weaknesses of the available evidence partially limit the applicability and generalisability of the findings.

**Clinical Trial Registration** PROSPERO registration number CRD42024485501 (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42024485501).

Extended author information available on the last page of the article

Published online: 22 March 2025

**Key Points**

The meta-meta-analysis showed that contrast therapy was most effective immediately post-intervention, massage therapy was most effective at 24 h, cryostimulation, kinesiotaping, contrast therapy, and compression were most effective at 48 h and kinesiotaping was most effective at 72 h post-intervention for reducing pain associated with delayed-onset muscle soreness.

The majority of the systematic reviews provided weak and unconvincing evidence, highlighting the need for larger higher-quality primary trials to better understand the effectiveness of various therapies for delayed-onset muscle soreness and to explore the potential synergistic effects of combined therapies.

Only 2 out of 29 systematic reviews were rated as high quality using the AMSTAR-2 (A MeaSurement Tool to Assess systematic Reviews-2) tool, while 17 out of 29 were of critically low quality. The most frequent methodological flaws included the lack of justification for excluded studies, unclear pre-established review methods and insufficient investigation of publication bias.

## 1 Introduction

Exercise-induced muscle damage pertains to micro-injuries during eccentric (lengthening) exercises and inflammatory processes in skeletal muscles, while subjective experiences such as muscle tightness and discomfort are typically referred to as delayed-onset muscle soreness (DOMS) [1]. Delayed-onset muscle soreness is characterised by muscle pain, stiffness and reduced function, which typically arise within 24–72 h post-exercise and subside by 5–7 days post-exercise [2]. The exact mechanisms underlying DOMS remain under research; however, it is believed that the loss of myofibrillar integrity with Z-band streaming and disruption of sarcomeres in the myofibrils lead to protein degradation, autophagy and the formation of pain sensation [3–5]. Damage to the neural microstructure may also play a large role in the formation of DOMS [6]. Delayed-onset muscle soreness induces a discharge of proinflammatory molecules including tumour necrosis factor-alpha, interleukin-1 beta, interleukin-6, interleukin-1 receptor antagonist, creatine kinase and lactate dehydrogenase into the blood [7]. Delayed-onset muscle soreness can negatively affect training sessions, decrease sport performance, and increase the risk of further injury if left untreated or mistreated [8, 9].

Various therapies have been promoted for treating or mitigating the effects of DOMS, with a range from pharmacological interventions [10] to nutritional strategies, not fully established [11, 12]. Physiotherapists (PTs) frequently use massage, stretching, vibration, photobiomodulation, low-intensity physical activity, compression, and kinesiotaping or apply heat/cold variously [13–17]. However, the evidence from systematic reviews (SRs) for the effectiveness of those therapies is often contradictory or inconclusive, further impeding the clinicians' decision making [13–16, 18, 19].

Differences in the conclusions of individual SRs, even those within the same therapy, may be due to search limitations and inclusion/exclusion of individual randomised controlled trials (RCTs). These inconsistencies can be addressed in an umbrella review (UR) by performing an overlap analysis, which helps to verify the comprehensiveness of publication coverage across the reviews [20, 21]. In addition, re-evaluating the included studies allows new more extensive meta-meta-analyses to be carried out than the previous more selective meta-analyses [22].

To the best of our knowledge, no UR exists on the topic as well as no study has comprehensively assessed the strength and quality of the scientific evidence on a multitude of therapeutic approaches for DOMS. Therefore, our study and mapping exercise has the potential to inform therapists, coaches, clinicians as well as athletes themselves on how to effectively reduce the symptoms of DOMS. Considering the increasing number of SRs of RCTs evaluating the wide variety of treatments of DOMS, it is reasonable to critically and collectively evaluate and map the scientific evidence for all those treatments in one UR [23]. We aimed to combine and analyse quantitative data from various systematic reviews and meta-analyses related to post-exercise physical therapy interventions for DOMS and graphically represent the existing evidence on the subject, categorising it based on the strength of evidence.

## 2 Methods

This is an UR (also referred as a systematic review of systematic reviews, or an overview of systematic reviews [21]) with a meta-meta-analysis, conducted according to the methodological criteria of such studies [24, 25]. By combining quantitative data from various meta-analyses addressing the same interventions and outcomes, meta-meta-analyses provide additional quantitative insights into the effects of interventions [26, 27]. Our UR also involved creating a map of the existing evidence and a methodological quality assessment of the available evidence. To avoid duplication and research waste, and to ensure transparency, the study was registered at PROSPERO (CRD42024485501) and the

detailed protocol was published open access [28]. We did not make any amendments to the initial protocol. We adhered to the Preferred Reporting Items for Overviews of Reviews (PRIOR) criteria while conducting and reporting this UR (see Electronic Supplementary Material [ESM]) [21].

### 2.1 Data Sources and Search Strategy

The following bibliographic databases and systematic reviews databases were searched for relevant SRs: Medical Literature Analysis and Retrieval System Online (MEDLINE) via Ovid, Excerpta Medica Database (Embase) via Ovid, Cochrane Database of Systematic Reviews (CDSR) by Wiley, Cumulative Index to Nursing and Allied Health Literature (CINAHL) by EBSCO, Epistemonikos and Physiotherapy Evidence Database (PEDro). Identified references were downloaded into the bibliographic management software for further handling (EndNote X9; Clarivate, Philadelphia, PA, USA). The search strategies were developed specifically for each database and the keywords were adapted to each database. Searches combined relevant search terms comprising indexed keywords (e.g. Medical Subject Headings and EMTREE) and free text terms appearing in the title and/or abstract of database records. Searches were limited by date, from January 1998 to January 2024, and not by language. The main MEDLINE strategy was independently peer reviewed by a second author (JZ). The peer review process was informed by the PRESS Peer Review of Electronic Search Strategies 2015 Guideline Statement [29]. References and citations in retrieved SRs were also checked. The complete search strategy for the databases can be found in the ESM.

### 2.2 Eligibility Criteria

We included SRs that fulfilled the following PICOST criteria:

*Population* SRs on healthy adults with DOMS excluding studies of medically compromised individuals. We included eligible participants of any level of physical activity or sports performance, with no restriction to the type of activity or sports discipline. We planned to exclude studies addressing the elderly (over 65 years of age). However, if an SR included a mixed population e.g. healthy or diseased individuals, or adults and other age groups, we extracted data on healthy adults only.

*Intervention and comparator* We included SRs of any type of PT interventions provided post-exercise, regardless of whether the signs or symptoms of DOMS occurred prior to the intervention. Preventive interventions (pre-exercise interventions) were considered ineligible. The interventions had to be administered, i.e. either applied (e.g. massage) or supervised (e.g. stretching) by PTs. Systematic review that did not report whether the interventions were implemented by PTs were considered eligible, based on the assumption that the interventions in question are within the competencies of PTs. We considered ineligible SRs on self-administered interventions, such as compression garments. The interventions could be compared against any comparators.

We excluded SRs of interventions other than by PTs, such as alternative and complementary therapies (e.g. acupuncture, reflexology, herbal medicine, homoeopathy), dietary supplements/nutritional interventions (e.g. amino acids, creatinine, beetroot, caffeine, curcumin, l-carnitine, omega-3 fatty acids, pomegranate, spirulina, vitamins C and E) or pharmacological interventions (e.g. cyclo-oxygenase 2 inhibitors, non-selective non-steroidal anti-inflammatory drugs), as well as preventive interventions (pre-exercise interventions) and self-administered interventions, such as compression garments.

*Outcome* The primary outcome was the intensity of post-exercise muscle soreness/pain at any endpoint, quantified with any measurement scale. We regarded any adverse effects (AEs) as secondary outcomes.

*Study type and timeframe* To be included, studies had to fulfil minimum methodological criteria for an SR, as defined by the Centre for Reviews and Dissemination guidance, i.e. providing sufficient information on reproducible search strategy, eligibility criteria and methodological quality/risk of bias (RoB) assessments [30]. Narrative/non-systematic reviews or those published before 1998 were excluded. Studies such as network meta-analyses, scoping reviews or otherwise labelled (e.g. 'evidence synthesis study') were considered for inclusion based on the above required criteria for an SR.

### 2.3 Study Selection

Titles and abstracts of records identified through electronic database searching were independently screened by two reviewers (PW and MC). During this initial phase of the screening process, any references that did not meet the inclusion criteria were excluded. Subsequently, full papers were obtained for all the remaining references. These were independently examined in detail by the same two reviewers to determine whether they met the criteria for inclusion in the review. Details of those studies assessed during full paper screening were reported in a table, including any reasons for exclusion from the review. Concerning both screening stages, any discrepancies and inconsistencies were resolved through discussion; with a third reviewer (SW) acting as an arbiter when necessary. We used the PRIOR flow diagram for the visual representation of the search and selection process [21].

### 2.4 Data Extraction

Data extraction sheets were individually designed and piloted in consultation with the research team, using Microsoft Excel® (version 2021; Microsoft Corporation, Redmond, WA, USA). The extracted data types include details of the populations, interventions (dose, frequency, intensity and duration), control groups, confounders and/or co-interventions, outcomes and effect estimates. We also extracted the date of the last database searches, number of RCTs included, total sample size, RoB in primary studies, whether meta-analysed (or not), the review authors' conclusions (and direction of conclusion, i.e. positive, negative, equivocal) and whether any AEs were reported. However, if an SR included a mixed population, e.g. healthy or diseased individuals, or adults and other age groups, we extracted data on healthy adults only. Data extraction was performed by two teams of reviewers independently (WP, PW and JZ, MC) with a third reviewer acting as arbiter in case of any disagreements (SW).

### 2.5 Methodological Quality Assessment

The methodological quality of the included SRs was assessed in duplicate by two reviewers (MC, JZ) using the AMSTAR-2 (A MeaSurement Tool to Assess Systematic Reviews) [31]. Any disagreements were resolved through discussion with supervisors (SW and MP). AMSTAR-2 can be applied to judge the methodological quality/RoB of systematic reviews including RCTs, non-RCTs or both [31]. AMSTAR-2 confidence in review ratings strongly correlates with the overall domain rating in another popular SR appraisal tool, the ROBIS [32]. The AMSTAR-2 tool consists of 16 questions. Each question can be scored as 'yes', 'no' or 'partial yes'. There are seven critical questions and nine non-critical questions. The overall quality assessment of an SR can fall into the following categories: 'high', 'moderate', 'low' or 'critically low' [31].

### 2.6 Evidence Map

In order to provide a graphical display of the evidence base of the subject matter, we created an evidence map using the following four dimensions, i.e. size and colour of bubbles, x-axis and y-axis, referring to study size, RoB, effect size and strength of evidence, respectively [23, 33]:

1. Number of cases: the size of each bubble corresponds directly to the number of cases in the experimental groups among studies included in the respective SRs after excluding overlapping RCTs.
2. Risk of bias: bubbles are colour coded, i.e. red indicating a very low percentage (0–15%) and blue indication a high percentage (40%) of studies at an overall low RoB assessed in the respective SRs (Jadad scale, PEDro scale, Cochrane RoB/RoB-2 tools).
3. Effect size ($x$-axis): therapies are categorised according to the effect size (standardised mean difference (SMD)/adjusted Hedges' $g$), only when the effect size favoured the intervention groups. When the effect favoured the controls or was not statistically significant ($p > 0.05$), it was classified as non-significant on the $y$-axis.
4. Strength of evidence ($y$-axis).

In this dimension, therapies are grouped into five personalised categories as described by [25]:

*Convincing (Class I)* when number of cases > 200, $p < 0.000001$, $I^2 < 25\%$, $p > 0.1$ Egger's test for small-study effects, $p > 0.1$ of a test for excess of significance bias (ESB) and a meta-analysis powered at least 80% to detect a SMD of 0.2.

*Highly suggestive (Class II)* when number of cases > 100, $p < 0.0001$, $I^2 \geq 25\%$ but <50%, $p > 0.05$ of an Egger's test for small-study effects, $p > 0.05$ of a test for ESB, meta-analysis has a power < 80% to detect a SMD of 0.2 but a power ≥ 80% to detect a SMD of 0.4 and Class I criteria not met.

*Suggestive (Class III)* when number of cases > 50, $p < 0.001$, $I^2 \geq 50\%$ but < 75%, $p > 0.01$ of an Egger's test for small-study effects, $p > 0.01$ of a test for ESB, meta-analysis has a power < 80% to detect a SMD of 0.4 but a power ≥ 80% to detect a SMD of 0.6 and Class I–II criteria not met.

*Weak (Class IV)* when $p < 0.05$ and class I–III criteria not met, and

Non-significant (Class V) when p > 0.05.

For the determination of the class, the fulfilment of all criteria was required. If any of the criteria were not met, the grade was lowered by one class [25].

### 2.7 Degree of Primary Study Overlap

We determined the degree of overlap by calculating the "covered area" (CA), using the formula CA = $N$/rc) and the "corrected covered area" (CCA), calculated by the formula CCA = $(N - r)/(r \cdot c - r)$ [20]. Here, "$N$" represents the total number of included publications, accounting for the double counting of overlapped trials. Additionally, "$r$" denotes the number of trials included, and "$c$" signifies the number of meta-analyses conducted. The interpretation of the CCA values provides insights into the extent of overlap: 0–5 indicates a slight overlap, 6–10 suggests a moderate overlap, 11–15 signifies a high overlap and values exceeding 15 suggest a very high overlap. Overlap analyses were conducted for each therapy independently within the identified SRs, irrespective of whether a meta-analysis was conducted. Following

a thorough examination of each paper included in the individual SRs, only duplicated RCTs within the Population, Intervention, Comparison, and Outcome (PICO) criteria defined in our UR were considered in the calculation [34]. We used the Graphical Representation of Overlap for OVErviews (GROOVE) tool [35] for a tabular and graphical (heat map) representation of overlap for each PICO.

## 2.8 Data Analysis and Synthesis

### 2.8.1 Qualitative Summary

Characteristics of the included SRs were summarised in a tabular form and narrative synthesis in terms of means and percentages. It includes a detailed overview of the characteristics and outcomes of each included SR. These findings were examined in light of the certainty of the recommendations made, based on the individual AMSTAR-2 assessments. This examination also involved identifying any potential factors that could introduce bias into the data or factors that might limit the credibility, reliability and generalisability of the findings.

### 2.8.2 Meta-meta Analysis

To estimate the effect size and stratify evidence according to the adapted criteria, we used the Metaumbrella package, version 1.0.9 for R (R Foundation for Statistical Computing, Vienna, Austria, https://www.rstudio.com [accessed 5 February 2024]) [36]. In instances where SRs did not report sufficient data by observation time and presented a combined effect, we extracted pertinent data, i.e. means, standard deviations and 95% confidence intervals (CIs). If no figures were available, we extracted the data from the available graphs using WebPlotDigitizer software (Version 4.6; WebPlotDigitizer, Pacifica, CA, USA). Extracting data from figures is faster, more reliable and reduces dependency on authors compared to requesting precise values directly from them [68].

We used the restricted likelihood maximum estimator to quantify the between-study variance in the random-effects meta-analysis. We assessed the significance of pooled SMDs and 95% CIs with adjusted Hedges' $g$ to address the potential overestimation of the true population effect size in smaller studies [37, 38]. The effect size categorisation was set as: 0–0.19 = negligible effect, 0.20–0.49 = small effect, 0.50–0.79 = moderate effect and ≥ 0.80 = large effect [39, 40]. Heterogeneity was assessed using the $I^2$ statistic with values ≥ 75% indicating high, > 50% moderate and > 25% low heterogeneity, respectively [41]. Funnel plot asymmetry (small study effects) was evaluated with Egger's test [42].

Finally, to measure whether there is an excess of studies with statistically significant results (ESB) we used a combined method (TESSPSST) of Test of Excess Statistical Significance (TESS) and the Proportion of Statistical Significance Test (PSST) [43]. Excess significance was considered at $p < 0.10$.

### 2.8.3 Subgroup Analyses

We performed subgroup analyses by specific therapies, e.g. cooling therapy or cryostimulation or type of manual/massage therapy. In the context of this review, *cooling therapy* refers to general cooling techniques such as the application of ice packs, cold compresses or cold water immersion. In contrast, *cryostimulation* encompasses more advanced cooling techniques, such as whole-body cryostimulation or localised cryostimulation, where extremely low temperatures (below − 100 °C) are applied using cold air or liquid nitrogen chambers. However, incomplete information provided in individual SRs and varying reporting methods prevented us from performing some of the initially planned subgroup analyses such as intensity or duration or frequency of interventions, athletic discipline, sex, age, physical activity level, RoB and type of control [28]. Based on the SRs, it was possible to conduct cross-sectional analyses, and these took into account, for individual therapies, the characteristics of age, group characteristics by sex, duration or frequency of intervention. If data extracted from SRs were doubtful (incomplete or unclear), we reanalysed directly individual RCTs included in a given SR.

# 3 Results

## 3.1 Study Selection

The electronic searches generated a total of 265 records. After deduplication, titles and abstracts of 182 records identified via database searching and four via hand searching of reference lists and citation tracking were screened against eligibility criteria. Finally, 29 SRs met the inclusion criteria. The PRIOR diagram of the search and selection process is presented in Fig. 1. The list of included SRs and excluded papers, with reasons for their exclusion, are listed in the ESM.

## 3.2 Characteristics of the Included Systematic Reviews

Eligible SRs were published between 2003 and 2023. Four, out of 29 included SRs, [13, 44–46] did not employ meta-analytic techniques, and one reported a network meta-analysis [17]. Included SRs evaluated 863 unique RCTs, with a total of 25,523 participants, mainly young adults (age range 15–64 years). The populations were relatively

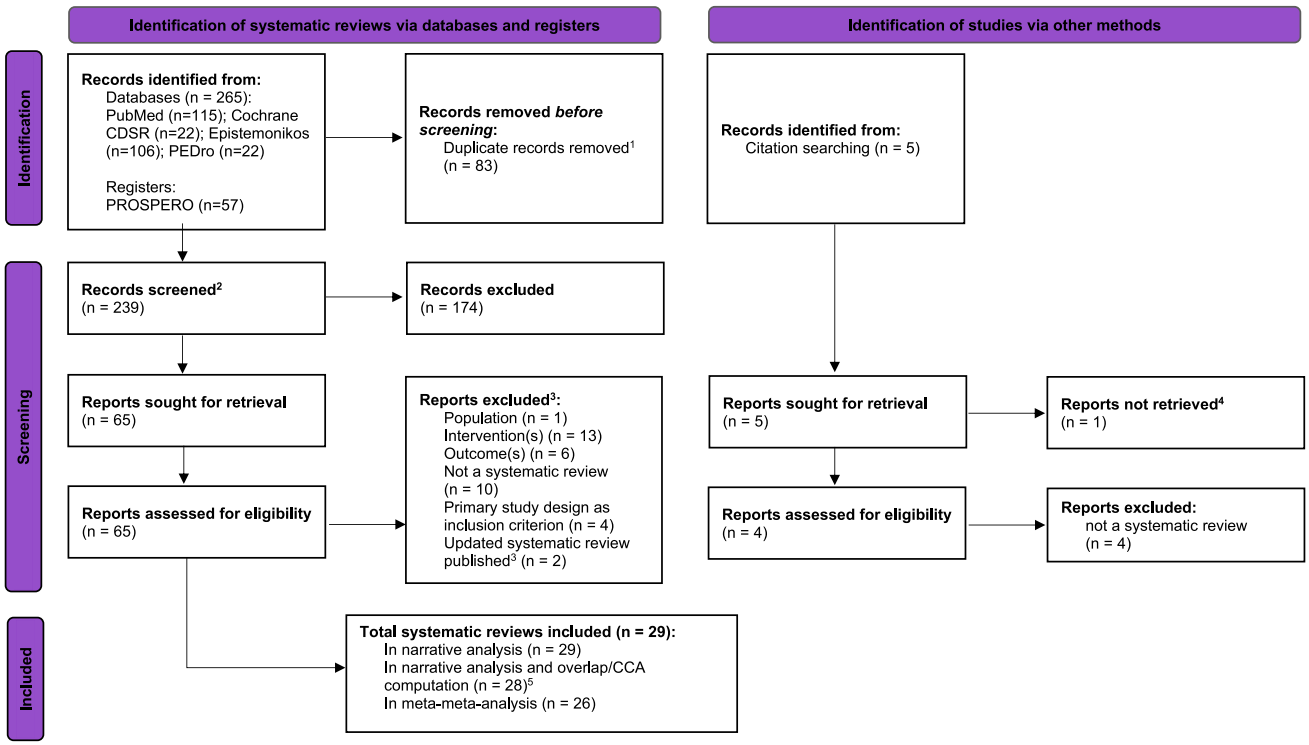


**Fig. 1** Preferred Reporting Items for Overviews of Reviews (PRIOR) flow diagram of the systematic review search and selection process. Note: As we did not search for supplemental primary studies, we used the simplified PRIOR flow diagram [21]. [1] Sixty-one records removed using the Endnote filter and additional 22 duplicates removed identified through searches in the PEDro database; [2] Potentially eligible records identified in the PROSPERO database search, whose final reports were not published (i.e. neither reported at PROSPERO as "review completed and published" nor located via database searching), ongoing or discontinued reports registered at PROSPERO are listed in the ESM, 9 records; remaining PROSPERO records were excluded based on title screening; [3] Several reports excluded based on more than one reason (e.g. ineligible intervention and design of included primary studies, see ESM for details), but one primary reason for exclusion reported in the flow chart (as per PRIOR template); [4] One report [14] did not include a list of primary studies included for analysis, and the authors did not respond to requests; [5] Decision made based on full text and overlap analyses (not on title and/or citation data analyses); please see the ESM for details. *CCA* corrected covered area, *CDSR* Cochrane Database of Systematic Reviews

homogeneous and included healthy (mainly active) individuals. For detailed data on individual therapies (based on a subgroup analysis of SRs containing meta-analyses), please see the ESM.

The interventions were, however, heterogeneous and included (as labelled in the included SRs): acupuncture, [13, 16] blood flow restriction, [45, 46] contrast water therapy, [14, 16–18, 47–49] cold water immersion, [15–18, 48–54] compression therapy, [13, 16, 55] cryostimulation, [14, 16–18, 49, 52–54, 56] electrical stimulation, [13, 14, 16, 57, 58] high voltage pulsed current [16, 58]), kinesiotaping, [16, 59] phototherapy (light-emitting diode therapy [16, 50, 60], photobiomodulation [16, 50], low level laser therapy [16, 60]), low-intensity exercise [13–16, 45], massage [13–16, 19, 61], magnetic therapy [16], electrical stimulation (microcurrent electrical neuromuscular stimulation [13, 16, 58], neuromuscular electrical stimulation [16, 57], interferential current [13, 16, 58], transcutaneous electrical nerve stimulation [13, 16, 58], shortwave diathermy [16], stretching [1, 13, 15, 16, 44, 62], ultrasound [13, 16, 17, 54], vibration therapy [63] and whole body cryostimulation [52, 56] Control groups ranged from passive recovery or rest, placebo or sham therapy, compression therapy, cold application, proprioceptive neuromuscular facilitation, partial body cryostimulation, relaxation biofeedback, stretching, transcutaneous electrical nerve stimulation or usual care. There was a considerable range of outcome measures. Twenty-six SRs did not report data regarding AEs and four reported that no AEs were identified in the included primary studies. Thirteen (43%) of the included SRs were funded by national organisations or public grants, one (3%) was sponsored by public and industry funds, ten SRs (34%) declared no funding, while five SRs (17%) failed to provide information on funding. Systematic reviews analysed the RoB or methodological quality of included studies: 62% used the Cochrane RoB tool (or its second version, RoB-2), 34% used the PEDro scale and one study used the Jadad scale [44] Only three out of 29 included SRs (10.3%) provided lists of

excluded studies, with justifications—these were Cochrane reviews [48, 56, 62] A narrative summary of the included SRs is provided in Table 1, while the ESM provides additional characteristics of the included SRs (such as muscle groups investigated), respectively.

## 3.3 Methodological Quality of the Included Systematic Reviews

Only two SRs, out of 29 (7%) were assessed as high quality with the AMSTAR-2 tool [1, 48] while 17 studies (59%) were of critically low quality [13, 15–17, 44–47, 49, 51, 53, 55, 57, 61, 63, 64] Nine studies (34%) were of low quality [14, 18, 19, 50, 52, 56, 58–60] and one SR (3%) was of moderate quality [62]

As per the AMSTAR-2 critical items, the lack of a list of excluded studies with a justification was the most common methodological flaw (26 studies, 89.7%), and 12 studies (41%) lacked a clear statement that review methods were established before the conduct of the review. The third most common flaw was the lack of an adequate investigation of publication bias. However, 28 (97%) of the studies used an appropriate RoB tool for primary studies and only one study failed to do so. Among SRs using meta-analytic techniques, the majority used appropriate methods for pooling and similarly most addressed the results of the RoB assessments while interpreting and/or discussing their results. We should highlight that full details of search strategies were frequently missing. The distribution of AMSTAR-2 items assessment is shown in Fig. 2, while detailed AMSTAR-2 assessments are provided in the ESM.

## 3.4 Pain Reduction in DOMS; Meta-meta-analysis of Therapeutic Interventions

The effect size (Hedges' $g$) and publication bias (Egger's test, ESB test) for each intervention and follow-up time after the DOMS induction varied widely (as reported in the ESM). Figure 3 reports significant meta-analysis results grouped by factor (therapy) and time of the outcome (pain) reported. Interventions ineffective in reducing pain/soreness (non-significant Hedges' $g$) for the given timepoints are shown in the ESM.

## 3.5 Evidence Map and Strength of Evidence

Below, we present the results of our reanalysis of the primary data included in each meta-analysis. They summarise the evidence on the effects of the interventions considered in the included SRs, specifically regarding the timing after their application.

### 3.5.1 Immediately Post-Intervention

With the highest strength of evidence, contrast therapy showed a significant effect in reducing pain immediately post-intervention (Class II, high RoB in primary studies, $g = 0.67$ [95% CI 0.95, 0.38]), whereas cooling therapy (moderate RoB in primary studies, $g = 0.55$ [95% CI 0.77, 0.33]) and cryostimulation (high RoB in primary studies, $g = 0.70$ [95% CI 1.30, 0.09]) showed a significant moderate effect size but a lower strength of evidence (Class IV) [Fig. 3]. Massage, heat therapy and exercise did not reduce DOMS significantly (Class V) [Fig. 4].

### 3.5.2 24 Hours Post-Intervention

With the highest strength of evidence, massage therapy showed a significant effect in reducing pain at 24 h post-intervention (Class III, high RoB in primary studies, $g = 0.41$ [95% CI 0.71, 0.10]), whereas cooling therapy (moderate RoB in primary studies, $g = 0.48$ [95% CI 0.66, 0.30]), contrast therapy (high RoB in primary studies, $g = 0.48$ [95% CI 0.86, 0.10]), electrical stimulation (moderate RoB in primary studies, $g = 0.57$ [95% CI 1.05, 0.08]), cryostimulation (high RoB in primary studies, $g = 0.76$ [95% CI 1.40, 0.12]), phototherapy (low RoB in primary studies, $g = 0.84$ [95% CI 1.55, 0.12]) and heat therapy (high RoB in primary studies, $g = 1.64$ [95% CI 3.10, 0.18]) showed significant low to large effect sizes but lower strength of evidence (Class IV) (Fig. 4). Compressions, vibration therapy, kinesiotaping, stretching and exercise did not significantly reduce DOMS (Class V) (Fig. 5).

### 3.5.3 48 Hours Post-Intervention

Compression ($g = 0.52$ [95% CI 0.82, 0.21]), contrast therapy ($g = 0.57$ [95% CI 0.94, 0.19]), kinesiotaping ($g = 0.68$ [95% CI 1.08, 0.29]) and cryostimulation showed a significant effect in reducing pain at 48 h post-intervention and achieved the highest strength of evidence (Class III, high RoB in primary studies), whereas cooling therapy (moderate RoB in primary studies, $g = 0.36$ [95% CI 0.60, 0.12]), massage (high RoB in primary studies, $g = 1.12$ [95% CI 1.77, 0.46]), phototherapy (low RoB in primary studies, $g = 1.50$ [95% CI 2.08, 0.92]) and heat therapy (high RoB in primary studies, $g = 1.82$ [95% CI 3.18, 0.46]) showed significant low to large effect sizes but a lower strength of evidence (Class IV) (Fig. 5). Electrical stimulation, vibration therapy, stretching and exercise did not significantly reduce DOMS (Class V) (Fig. 6).

**Table 1** Characteristics of the included systematic reviews. Overview of the included systematic reviews in regard to the PICOs of interest

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Afonso et al., 2021 [1] | 5/111 (11/289) | Male and female individuals, 18–27 years old, athletes (jiu-jitsu fighters), non-athletes (sedentary) | Stretching (static passive, static active, dynamic, PNF) | Passive recovery, rest, massage | Pain (VAS, PPT) | High, $n=5$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB 2.0 tool) | 24 h: $g=1.55$; 95% CI 1.12, 1.97; $p<0.001$; $I^2=48.3\%$; Egger's test $p=0.231$<br>48 h: $g=1.50$; 95% CI 1.02, 1.98; $p<0.001$; $I^2=59.8\%$; Egger's test $p=0.257$<br>72 h: $g=0.98$; 95% CI 0.67–1.28; $p<0.001$; $I^2=0.0\%$; Egger's test $p=0.525$ | High |
| Apostolopoulos et al., 2015 [44] | 2/73 (152/5811) | Sex (NR), age (NR), non-athletes | Stretching (dynamic, static, active, PNF, passive, static ballistic, dynamic) | NR | Pain (numerical, visual, PPT) | High quality, $n=0$<br>low quality, $n=2$ (Jadad scale) | No meta-analysis conducted | Critically low |
| Azevedo et al., 2022 [52] | 5/95 (6/120) | Sex (NR), 22–39 years old, athletes (middle- and long-distance runners), physically active, well-trained individuals | Whole body cryotherapy (cryostimulation or cooling therapy) (10 to − 140 °C, 3–10 min), partial body cryotherapy (cryostimulation or cooling therapy) (15 to − 135 °C, 30–600 secs) | PBC (8 to − 135 °C, 30 s to 12 min) | Pain (VAS) | 4.4 ± 0.5 (PEDro scale) | 24 h: SMD = − 0.25; 95% CI − 0.66, 0.15; $p=0.23$; $I^2=0\%$ 48 h: SMD = − 0.08; 95% CI − 0.48, 0.33; $p=0.72$; $I^2=0\%$ | Low |
| Batista et al., 2023 [18] | 44/801 (44/801) | Male and female individuals, 14–41 years old, athletes (football players, jiu-jitsu competitors, netball players, hockey players, MMA athletes, bicycle motocross athletes, rugby players, basketball players, cyclists, triathletes, endurance runners), non-athletes, recreationally active | Cold water immersion (5–15 °C, 10–15 min intermittent or continuous application) | Passive recovery | Pain (VAS) | 4.6 ± 0.9 (PEDro scale) | Immediate effect, severe cold 5–9 °C: MD = 0.99; 95% CI 0.38, 1.60; $p=0.001$; $I^2=54\%$<br>Immediate effect, moderate cold 10–15 °C: MD = 1.00; 95% CI 0.44, 1.56; $p<0.001$; $I^2=87\%$<br>Delayed effect severe cold 5–9 °C: MD = 0.68; 95% CI 0.09, 1.26; $p=0.02$; $I^2=37\%$<br>Delayed effect moderate cold 10–15 °C: MD = 0.68; 95% CI 0.09, 1.26; $p=0.02$; $I^2=37\%$ | Low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Bieuzen et al., 2013 [47] | 13/305 (18/356) | Male and female individuals, 14–27 years old, athletes (football players, rugby players, trained netball players, junior elite rowers, strength trained, endurance-trained), non-athletes | Contrast water therapy (8–45 °C, 6–96 min, 1–4 sessions) | Passive recovery | Pain (VAS, Likert scale, PPT) | High, $n=13$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | < 6 h: SMD = − 0.62; 95% CI − 0.95, − 0.28; $p=0.0003$; $I^2=0\%$<br>24 h: SMD = − 0.51; 95% CI − 0.75, − 0.27; $p=0.0002$; $I^2=67\%$<br>48 h: SMD = − 0.58; 95% CI − 0.85, − 0.31; $p<0.00001$; $I^2=56\%$<br>72 h: SMD = − 0.40; 95% CI − 0.76, − 0.03; $p=0.03$; $I^2=54\%$<br>96 h: (SMD = − 1.21; 95% CI − 2.03, − 0.39; $p=0.004$; $I^2=$NA | Critically low |
| Bleakley et al., 2012 [48] | 12/251 (17/366) | Male and female individuals, 19–35 years old, athletes (cyclists, netballers, basketball players, soccer players), non-athletes | Cold water immersion (5–15 °C, 5–24 min; 3–5 × 1 min, 2 × 5 min or 2 × 15 min immersion | Passive recovery | Pain (VAS, Likert scale) | High, $n=12$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | Immediately: SMD = − 0.07; 95% CI − 0.43, 0.28; $p=0.68$; $I^2=2\%$<br>24 h: SMD = − 0.55; 95% CI − 0.84, − 0.27; $p=0.0001$; $I^2=64\%$<br>48 h: SMD = − 0.66; 95% CI − 0.97, − 0.35; $p<0.0001$; $I^2=57\%$<br>72 h: SMD = − 0.93; 95% CI − 1.36, − 0.51; $p<0.0001$; $I^2=80\%$<br>96 h: SMD = − 0.58; 95% CI − 1.00, − 0.16; $p=0.006$; $I^2=46\%$ | High |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Costello et al., 2015 [56] | 4/84 (4/84) | Male and female individuals, 20–31 years old, athletes (runners), physically active, non-athletes | WBC (− 110 to − 190 °C, 3 min) | Passive recovery, no treatment or placebo | Pain (VAS) | High, $n=4$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | 1 h: SMD = − 0.77; 95% CI − 1.42, − 0.12; $p=0.02$; $I^2=0\%$<br>24 h: SMD = − 0.57; 95% CI − 1.48, 0.33; $p=0.21$; $I^2=64\%$<br>48 h: SMD = − 0.58; 95% CI − 1.37, 0.21; $p=0.15$; $I^2=53\%$<br>72 h: SMD = − 0.65; 95% CI − 2.54, 1.24; $p=0.50$; $I^2=87\%$<br>96 h: SMD = − 0.33; 95% CI − 0.95, 0.30; $p=0.31$; $I^2=0\%$<br>120 h: SMD = − 0.32; 95% CI − 1.16, 0.52; $p=0.46$; $I^2=$ NA% | Low |
| Davis et al., 2020 [61] | 10/311 (29/1012) | Male and female individuals, age (NR), athletes (amateur boxers, recreational runners, collegiate basketball and volleyball players, triathletes) | Massage (20–30 min) | Passive recovery, detuned ultrasound | Pain (NR) | 4.7 ± 1.6 (PEDro Scale) | Overall effect: SMD = 1.13; 95% CI 0.44, 1.82; $p<0.05$; $I^2=86\%$ | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Dupuy et al., 2018 [14] | 80/1188 (99/1454) | Male and female individuals, age (NR), competitive athletes (NR), recreational athletes, inactive people | Active recovery (NR), massage (20–30 min), NMES (NR), stretching (NR), cold water immersion (11–15 °C over 11–15 min), contrast water therapy (NR), cryotherapy-cryostimulation (NR) | Passive recovery | Pain (NR) | PEDro scale and van Tulder scale (NR) | Overall effect:<br>Active recovery: SMD = −0.94; 95% CI −1.61, −0.28; $I^2$ = NA<br>Stretching: SMD = 0.15; 95% CI 0.00, 0.29; $I^2$ = NA<br>Massage: SMD = −2.26; 95% CI −3.05, −1.47; $I^2$ = NA<br>Compression garments: SMD = −0.92; 95% CI −1.34, −0.50; $I^2$ = NA<br>Electrostimulation: SMD = −0.28; 95% CI −0.59, 0.03; $I^2$ = NA<br>Immersion: SMD = −0.47; 95% CI −0.77, −0.18; $I^2$ = NA<br>CWT: SMD = −0.40; 95% CI −0.73, −0.07; $I^2$ = NA<br>Cryotherapy: SMD = −0.53; 95% CI −1.04, −0.03; $I^2$ = NA<br>Hyperbaric therapy: SMD = 0.55; 95% CI −0.12, 1.22; $I^2$ = NA | Low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Ferlito et al., 2022 [50] | 3/50 (4/66) | Male individuals, mean 24.6 ± 4.9 years old, athletes (futsal players), non-athletes | Photobiomodulation (30–300 secs, continuous output and 0–1500 Hz) | Cryotherapy-cooling therapy (thermal bags or cold-water immersion, 5 – 20 min) | Pain (VAS) | High, $n = 0$<br>Low, $n = 1$<br>Unclear, $n = 2$ (Cochrane RoB tool) | Immediately: SMD = 1.21; 95% CI − 0.43, 2.86; $p = 0.15$; $I^2 = 79\%$<br>1 h: SMD = 1.61; 95% CI = 0.84, 2.39; $p < 0.0001$; $I^2 = 0\%$<br>24 h: SMD = 1.92; 95% CI = 0.61, 3.23; $p = 0.004$; $I^2 = 59\%$<br>48 h: SMD = 2.25; 95% CI 1.13, 3.37; $p < 0.0001$; $I^2 = 37\%$<br>72 h: SMD = 1.91; 95% CI 1.08, 2.73; $p < 0.00001$; $I^2 = 0\%$<br>96 h: SMD = 1.70; 95% CI 0.65, 2.76; $p = 0.002$; $I^2 = NA$ | Low |
| Guo et al., 2017 [19] | 11/239 (11/239) | Male and female individuals, 19–41 years old, athletes (volleyball, basketball players, bodybuilders), non-athletes | Massage (effleurage and petrissage, pushing, swing, grasping, vibrating and plucking; 6–30 min) | Passive recovery, unilateral shoulder shrugs | Pain (VAS, PPT) | High, $n = 11$<br>Unclear, $n = 0$<br>Low, $n = 0$ (Cochrane RoB tool) | 24 h: SMD − 0.61; 95% CI − 1.17, − 0.05; $p = 0.03$; $I^2 = 87\%$<br>48 h: SMD = − 1.51; 95% CI − 2.24, − 0.77; $p < 0.0001$; $I^2 = 82\%$<br>72 h: SMD = − 1.46; 95% CI − 2.59, − 0.33; $p = 0.01$; $I^2 = 82\%$ | Low |
| Herbert et al., 2011 [62] | 5/158 (9/2551) | Male and female individuals, 18–40 years old, non-athletes | Stretching (static, ballistic, PNF 5–12 sessions and for 12 weeks, 40–900 secs) | No intervention, placebo pills, passive recovery | Pain (VAS, OS) | High, $n = 5$<br>Unclear, $n = 0$<br>Low, $n = 0$ (Cochrane RoB tool) | 18–30 h: MD = − 1.04; 95% CI = − 6.88, 4.79; $p = 0.73$; $I^2 = 0\%$<br>42–54 h: MD = 1.12; 95% CI − 4.63, 6.87; $p = 0.70$; $I^2 = 0\%$<br>66–78 h: MD = − 0.03; 95% CI − 7.49, 7.43; $p = 0.99$; $I^2 = 0\%$ | Moderate |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Hohenauer et al., 2015 [49] | 27/NR (36/574) | Sex (NR), 18–31 years old, athletes (junior soccer players, basketball players, football players, climbers, netball players), non-athletes | Cold water immersion (3–15 °C) WBC (− 10 °C, − 60 °C and − 110 °C), cold air application (− 30 °C), ice cuffs, cold packs application, ice bags application | Passive recovery (sitting, standing or resting supine-room temperature 15–24 °C), thermoneutral immersion (35 °C) | Pain (VAS, Borg, Likert scale) | High, $n=27$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | 24 h: $g=-0.69$; 95% CI − 1.06, − 0.32; $p<0.001$; $I^2=78.4\%$<br>48 h: $g=-0.62$; 95% CI − 1.00, − 0.25; $p<0.001$; $I^2=73.3\%$<br>72 h: $g=-0.18$; 95% CI − 0.53, 0.17; $p=0.013$; $I^2=54.3\%$<br>96 h: $g=-0.65$; 95% CI − 1.00, − 0.30; $p=0.238$; $I^2=24\%$ | Critically low |
| Lin et al., 2021 [59] | 8/289 (8/289) | Male and female individuals, 18–43 years, athletes (NR), non-athletes | Kinesiotaping (stretched 110–140%, worn 2–5 days) | Passive recovery, placebo | Pain (VAS) | High, $n=8$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | 24 h: MD = − 0.39; 95% CI − 1.14, 0.36; $p=0.31$; $I^2=71\%$<br>48 h: MD = − 0.67; 95% CI − 1.10, 0.24; $p=0.002$; $I^2=61\%$<br>72 h: MD = − 0.81; 95% CI − 1.45, − 0.17; $p=0.01$; $I^2=86\%$ | Low |
| Lu et al., 2019 [63] | 8/240 (10/302) | Male and female individuals, age (NR), athletes (soccer players), non-athletes, untrained | Vibration therapy (30 secs to 20 min, 12–120 Hz, 0.5 –5 mm amplitude) | Passive recovery, massage, static stretching | Pain (VAS, PPT) | High, $n=8$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | 24 h: SMD = − 1.53; 95% CI − 2.57, − 0.48; $p=0.004$; $I^2=91\%$<br>48 h: SMD = − 2.04; 95% CI − 3.40, − 0.69; $p=0.003$; $I^2=92\%$<br>72 h: SMD = − 1.60; 95% CI 2.99, − 0.21; = 0.02; $I^2=88\%$ | Critically low |
| Ma et al., 2020 [45] | 14/337 (23/599) | Male and female individuals, 18–35 years old, athletes (NR), non-athletes | Low-intensity exercise (10–60 min, training session, active exercise, run on the treadmill, downhill running, concentric exercise 40–70% intensity, 1–21 days, blood flow restriction (occlusion cuffs, 171–266 mmHg, 6–15 min, 2–3 cycles) | Passive recovery | Pain (soreness, PPT, 5-point pain scale) | 4.5 ± 1.8 (PEDro scale) | No meta-analysis conducted | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Malone et al., 2014 [57] | 13/189 (13/189) | Male and female individuals, 17–47 years old, athletes (taekwondo athletes, military, rock climbers, collegiate swimmers, soccer players, futsal players, baseball pitchers), non-athletes | NMES (1–8 Hz, 125–500 ms, 17,5–92 mA) | Passive recovery (lying supine, seating)<br>Active recovery (water exercise, cycling, running) | Pain (VAS, Ratings of perceptions of pain) | High, $n=1$<br>Unclear, $n=4$<br>Low, $n=8$ (Cochrane RoB tool) | No meta-analysis conducted | Critically low |
| Menezes et al., 2022 [58] | 13/349 (14/435) | Male and female individuals, 20–26 years old, athletes (amateur soccer players, strength athletes), non-athletes | MENS (0.3–30 Hz, 30–200 ms, 2–20 min), TENS HF/LF (4–110 Hz, 40–200 ms, 20–180 min), HVPC (125 Hz, 200 ms, 90 min), IFC (80–150 Hz, 125 ms, 30–300 min) | Placebo/sham (simulation or without electrostimulation)<br>Control (passive recovery) | Pain (VAS, PPT) | High, $n=6$<br>Unclear, $n=5$<br>Low, $n=2$ (Cochrane RoB tool) | TENS<br>Immediately: SMD = −0.39; 95% CI −1.24, 0.46; $p=0.37$; $I^2=0\%$<br>24 h: SMD = 0.11; 95% CI −0.36, 0.58; $p=0.65$; $I^2=0\%$<br>48 h: SMD = 0.12; 95% CI −0.49, 0.74; $p=0.70$; $I^2=48\%$<br>72 h: SMD = 0.20; 95% CI = −0.27, 0.68; $p=0.40$; $I^2=0\%$<br>96 h: SMD = 0.18; 95% CI −0.98, 1.34; $p=0.76$; $I^2=44\%$<br>TENS and HPVC<br>24 h: SMD = 0.15; 95% CI −0.51, 0.21; $p=0.42$; $I^2=0\%$<br>48 h: SMD = −0.04; 95% CI −0.45, 0.38; $p=0.85$; $I^2=0\%$<br>72 h: SMD = −0.06; 95% CI −0.52, 0.40; $p=0.80$; $I^2=0\%$ | Low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Nahon et al., 2021 [16] | 121/3373 (121/3373) | Male and female individuals, 13–45 years old, athletes (runners, MMA athletes, bodybuilders, BMX cyclists, soccer players, rugby players, basketball players), recreationally active, non-athletes | Magnetic therapy (clamps with magnetic disk, electromagnetic field applied 10 min, 5–7 days), contrast water therapy (1 min, 8–12 °C, 1–2 min 38–42 °C), vibration (3–30 min, 12–50 Hz, 2×7 days), cryotherapy- cooling therapy (12–30 min, -30–14 °C×3 days), cold water immersion (1–30 min, 5 °C ± 1–15 °C), Ice packs (15–60 min), partial body cryotherapy (3 min, − 110 °C), cold pack (15 min×5), microcurrents (0,3 Hz, 20 min for 3 days), ultrasounds (6–14 min, 0.16–1.3 W/cm2, 1–3 MHz), short wave diathermy (750 kHz, 10 min), phototherapy (6–12 min, 660–950 mm, 2,5–73 Hz, 3,2 –31.7 J), compression therapy (12–45 min, 10–220 mmHg), massage (30 secs per region to 30 min for 3 days), stretching (3–10 sets, 20–60 s, 10–30 s rest, 3×day), active exercise (10 series×10 repetitions – active movements, 15 min—submaximal jogging, 5 min – exercise bike), kinesiotaping (3–7 days, 30% stretch), acupuncture (10–15 min) | Passive recovery, no intervention | Pain (VAS, pressure algometry, McGill pain questionnaire, PPT) | 4.7 ± 1.5 (PEDro scale) | Magnetic therapy SMD = − 4.66; 95% CI − 11.22, 1.91; $p = 0.16$; $I^2 = 97\%$<br>Contrast therapy: SMD = − 1.05; 95% CI − 1.73, − 0.38; $p = 0.002$; $I^2 = 60\%$<br>Ultrasounds: SMD = − 2.89; 95% CI − 5.34, − 0.44; $p = 0.02$; $I^2 = 97\%$<br>Vibration: SMD = − 3.77; 95% CI − 6.35, − 1.19; $p = 0.004$; $I^2 = 96\%$<br>Phototherapy: SMD = − 6.77; 95% CI − 10.05, − 3.49; $p = 0.0001$; $I^2 = 95\%$<br>Electrical stimulation: SMD = − 3.97; 95% CI − 12.74, 4.80; $p = 0.37$; $I^2 = 100\%$<br>Cryotherapy: SMD = − 4.02; 95% CI − 6.58, − 1.46; $p = 0.002$; $I^2 = 100\%$<br>Stretching: SMD = 0.02; 95% CI − 1.32, 1.37; $p = 0.97$; $I^2 = 91\%$<br>Compression: SMD = − 2.33; 95% CI − 3.76, − 0.89; $p = 0.002$; $I^2 = 93\%$<br>Active exercise: SMD = − 2.03; 95% CI − 3.15, − 0.91; $p = 0.0004$; $I^2 = 93\%$<br>Massage: SMD = − 2.46; 95% CI − 3.52, − 1.41; $p = 0.00001$; $I^2 = 94\%$<br>Kinesiotaping: SMD = − 0.50; 95% CI − 1.18, 0.17; $p = 0.14$; $I^2 = 74\%$ | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Nampo et al., 2016 [60] | 4/140 (15/317) | Male and female individuals, 19–35 years old, athletes (volleyball and soccer players), non-athletes | Light-emitting diodes (630–880 nm, 10–500 mW, 36–208.5 J), laser therapy (660–950 nm, 30–534 mW, 4–380.4 J) | Placebo LED/LASER | Pain (VAS) | High, $n=4$<br>Unclear, $n=0$<br>Low, $n=0$ (Cochrane RoB tool) | 24 h: MD = − 3.81; 95% CI − 13.30, 5.68; $p=0.43$; $I^2=70\%$<br>48 h: SMD = − 8.13; 95% CI − 23.78, 7.52; $p=0.31$; $I^2=74\%$<br>72 h: SMD = − 9.80; 95% CI − 31.15, 11.56; $p=0.37$; $I^2=88\%$<br>96 h: SMD = − 5.15; 95% CI − 15.86, 5.55; $p=0.35$; $I^2=80\%$ | Low |
| Noquiera et al., 2020 [53] | 7/222 (7/222) | Male and female individuals, 18–29 years old, athletes (NR), recreationally active, non-athletes, untrained | Local cryotherapy-cooling therapy (20 min per day, 3 × 20 min per day), crushed ice application (15–60 min) | Passive recovery | Pain (Talag scale, VAS) | 7.3 ± 1.3 (PEDro scale) | 24 h: SMD = 0.44; 95% CI − 0.01, 0.89; $p=0.05$; $I^2=0\%$<br>48 h: SMD = − 0.36; 95% CI − 2.20, 1.48; $p=0.70$; $I^2=87\%$<br>72 h: SMD = − 0.46; 95% CI − 1.88, 0.95; $p=0.52$; $I^2=0.85\%$<br>96 h: SMD = − 0.07; 95% CI − 0.75, 0.62; $p=0.85$; $I^2=79\%$ | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| O'Connor and Hurley, 2003 [13] | 17/462 (30/883) | Male and female individuals, 18–40 years old, healthy subjects | Massage (8–30 min), stretching (4–10 min), cryotherapy-cooling therapy (ice massage 15–20 min, water immersion 5×20 min, 7×15 min), exercise (6×20 max contractions, 5×10 sub max concentric actions), warm whirlpool (38.9 °C×24 min), cold whirlpool (12.8 °C×24 min), ultrasounds (7–20 min, 0.8 –1.5 W/cm$^2$, 1 MHz), TENS (20 min, 90 ms, 80–90 pps), interferential currents (10–100 Hz, 30 min), microcurrent electrical stimulation (8 min, 0,3 Hz, 30 mA) | Passive recovery, upper body ergometer, light stretch, MES, placebo needling | Pain (Soreness Rating scale, VAS, GPRS) | 6.5 ± 3.5 (PEDro scale) | No meta-analysis conducted | Critically low |
| Rodrigues et al., 2022 [46] | 7/107 (8/118) | Male and female individuals, 18 – 26 years old, athletes (NR), recreationally active, non-athletes | Blood flow restriction (3–5 times 35– 220 mmHg) | Exercise without BFR | Pain (NPS, PPT) | 5.8 ± 0.9 (PEDro scale) | No meta-analysis conducted | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Torres et al., 2012 [15] | 34/790 (35/820) | Male and female individuals, 18–46 years old, athletes (NR), non-athletes | Cold water immersion 10–105 min, 5–15 °C), ice massage (15 min), massage (6–30 min), stretching (80–600 secs) | Without treatment, no intervention, contralateral leg or arm, light stretch, MES, upper body ergometer, thermoneutral water, stepping warmup, eccentric exercise | Pain (soreness, PPT, muscle stiffness) | 5.5 ± 3.5 (PEDro scale) | Massage<br>1 h: SMD = −0.11; 95% CI −0.39, 0.18; $p = 0.46$; $I^2 = 0\%$<br>24 h: SMD = −0.33; 95% CI −0.59, −0.07; $p = 0.01$; $I^2 = 0\%$<br>48 h: SMD = −0.96; 95% CI −2.02, 0.09; $p = 0.07$; $I^2 = 46\%$<br>72 h: SMD = 0.26; 95% CI −0.01, 0.56; $p = 0.06$; $I^2 = 0\%$<br>Stretching:<br>24 h: SMD = 0.22; 95% CI −0.36, 0.79; $p = 0.46$; $I^2 = 38\%$<br>48 h: SMD = 0.06; 95% CI −0.37, 0.48; $p = 0.79$; $I^2 = 0\%$<br>72 h: SMD = 0.14; 95% CI −0.28, 0.57; $p = 0.51$; $I^2 = 0\%$<br>LIE:<br>1 h: SMD = −0.75; 95% CI −1.85, 0.36; $p = 0.18$; $I^2 = 0\%$<br>24 h: SMD = −0.44; 95% CI −2.40, 1.51; $p = 0.66$; $I^2 = 98\%$<br>48 h: SMD = −0.64; 95% CI −2.65, 1.37; $p = 0.53$; $I^2 = 98\%$<br>72 h: SMD = −0.85; 95% CI −1.71, 0.01; $p = 0.05$; $I^2 = 92\%$ | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Wang et al., 2021 [54] | 32/1098 (32/1098) | Male and female individuals, 16–79 years old, athletes (NR), non-athletes | Cold water immersion (2.5 –25 min, 3–15 °C), cryotherapy-cryostimulation (3–4 min, room temperature − 10 to − 110 °C), cold pack (20–60 min, − 18 to 1.7 ± 2.8 °C)<br>hot pack (20–480 min, 43.3 ± 2.8 °C), air pulsed cryotherapy (5 min × 4, 5–15 °C), ice massage (15 min, ice ball), hot water immersion (14 min, 38 °C)<br>warm whirlpool (24 min, 38.9 °C)<br>infrared radiation (30 min)<br>sauna (15 min, 170–180 °C)<br>phase change material (3 h, 15 °C)<br>low-dose ultrasounds (10 min × 5 days) | Passive recovery, no intervention | Pain (VAS, muscle soreness questionnaire, Likert scale, modified Talag scale) | High, $n = 4$<br>Unclear, $n = 20$<br>Low, $n = 8$ (Cochrane RoB tool) | Colt therapy:<br>< 24 h: SMD = − 0.57; 95% CI − 0.89, − 0.25; $p = 0.0005$; $I^2 = 66\%$<br>> 24 h: SMD = − 0.37; 95% CI − 0.75, 0.00; $p = 0.05$; $I^2 = 75\%$<br>Heat therapy:<br>< 24 h: SMD = − 1.17; 95% CI − 2.28, − 0.09; $p = 0.03$; $I^2 = 91\%$<br>> 24 h: SMD = − 0.82; 95% CI − 1.38, − 0.26; $p = 0.004$; $I^2 = 82\%$ | Critically low |
| Wang et al., 2022 [17] | 59/1367 (59/1367) | Male and female individuals, 18–41 years old, athletes (NR), non-athletes | Cold water immersion, phase change material, cryotherapy-cryostimulation, contrast water therapy, hot/warm-water immersion, cold pack, hot pack, ice massage, ultrasounds | Passive recovery, no intervention, placebo, warm water immersion, WBC, PBC | Pain (VAS, GPRS, Likert scale, modified Talag scale) | High, $n = 13$<br>Unclear, $n = 35$<br>Low, $n = 11$ (Cochrane RoB tool) | Network MA | Critically low |
| Wiśniowski et al., 2022 [55] | 9/262 (12/322) | Male and female individuals, 18–56 years old, athletes (endurance runners, strength trained, wheelchair basketball and rugby players, resistance-trained), non-athletes | Pressotherapy (60–260 mmHg, 20–60 min) | Passive recovery, placebo therapy | Pain (VAS, 5-point pain rating scale, overall fatigued) | High, $n = 5$<br>Unclear, $n = 4$<br>Low, $n = 0$ (Cochrane RoB 2 tool) | 24 h: SMD = − 0.28; 95% CI − 0.60, 0.04; $p = 0.09$; $I^2 = 43\%$<br>48 h: SMD = − 0.40; 95% CI − 0.73, − 0.07; $p = 0.02$; $I^2 = 0\%$<br>72 h: SMD = − 0.37; 95% CI − 0.79, 0.05; $p = 0.08$; $I^2 = 24\%$<br>96 h: SMD = − 0.36; 95% CI − 0.72, 0.00; $p = 0.05$; $I^2 = 0\%$ | Critically low |

**Table 1** (continued)

| Systematic review, year [ref] | RCTs/participants in SR under PICO criteria (total RCTs/participants in SR) | Population (sex, age, sport practised) | Intervention (dose, frequency, intensity, duration) | Control/comparison intervention(s) | Outcome measure (s) | Methodological quality/risk of bias in primary studies: authors' assessment (scores/assessment tool), *n* RCTs | Intervention vs control/comparator (SMD/MD/*g*, 95% CI, heterogeneity) | AMSTAR-2 quality assessment |
|---|---|---|---|---|---|---|---|---|
| Xiao et al., 2023 [51] | 7/218<br>20/419 | Male and female individuals, 18–40 years old, athletes (runners, swimmers, cyclists, soccer players, rugby players, endurance cyclists), non-athletes | Cold water immersion (5–15 °C to the iliac crest, 6–15 °C to the midsternal level and shoulders) | Passive recovery (sitting in a room with a temperature from 6–32 °C, and humidity from 37–62%) | Pain (VAS) | High, $n=5$<br>Unclear, $n=0$<br>Low, $n=2$ (Cochrane RoB tool) | Immediately: SMD = − 0.59; 95% CI − 0.90, − 0.28; $p=0.0002$; $I^2=1\%$<br>24 h: SMD = − 0.34; 95% CI − 0.65, − 0.04; $p=0.03$; $I^2=67\%$<br>48 h: SMD = − 0.25; 95% CI − 0.58, 0.07; $p=0.12$; $I^2=66\%$ | Critically low |
| Yin et al., 2022 [64] | 13/355<br>16/431 | Sex (NR)<br>19–36 years old, athletes (NR) | Vibration therapy (3–120 Hz, 1–20 min, 1.5–5 mm amplitude) | Passive recovery, stretching or ultrasound | Pain (VAS, PPT) | High, $n=0$<br>Unclear, $n=10$<br>Low, $n=3$ (Cochrane RoB tool) | 24 h: SMD = − 1.10; 95% CI − 2.04, − 0.32; $p=0.007$; $I^2=90\%$<br>48 h: SMD = − 2.55; 95% CI − 3.93, − 1.18; $p=0.0003$; $I^2=94\%$<br>SMD = − 1.40; 95% CI − 2.28, − 0.52; $p=0.002$; $I^2=95\%$ | Critically low |

Nnumber of RCTs/total *N* = RCTs also denotes qRCTs, cluster, as per the authors' inclusion criteria

Note: terminology and specific abbreviations are provided in the table as they appear in the analysed SRs (e.g. 'sedentary' or 'inactive', MVC or MVIC); NR, refers to the fact that the authors of the study did not provide any detailed information

*AMSTAR-2* A MeaSurement Tool to Assess systematic Reviews-2, *BFR* blood flow restriction, *GPRS* Graphic Pain Rating Scale, *h* hours, *HVPC* high-voltage pulsed current electrical stimulation, *IFC* interferential current, *LIE* low-intensity exercise, *MD* mean difference, *MENS* microcurrent electrical stimulation, *min* minutes, *NMES* neuromuscular electrical stimulation, *NPS* numeric pain scale, *OS* own/author's scale, *PBC* partial body cryostimulation, *PICO* Population, Intervention, Comparison, and Outcome, *PNF* proprioceptive neuromuscular facilitation, *PPT* pressure pain threshold, *RCTs* randomised controlled trials, *sec* seconds, *SMD* standardised mean difference, *SR* systematic review, *TENS* transcutaneous electrical nerve stimulation, *TENS HF* transcutaneous electrical nerve stimulation high frequency, *TENS LF* transcutaneous electrical nerve stimulation low frequency, *VAS* visual analogue scale, *WBC* whole body cryostimulation

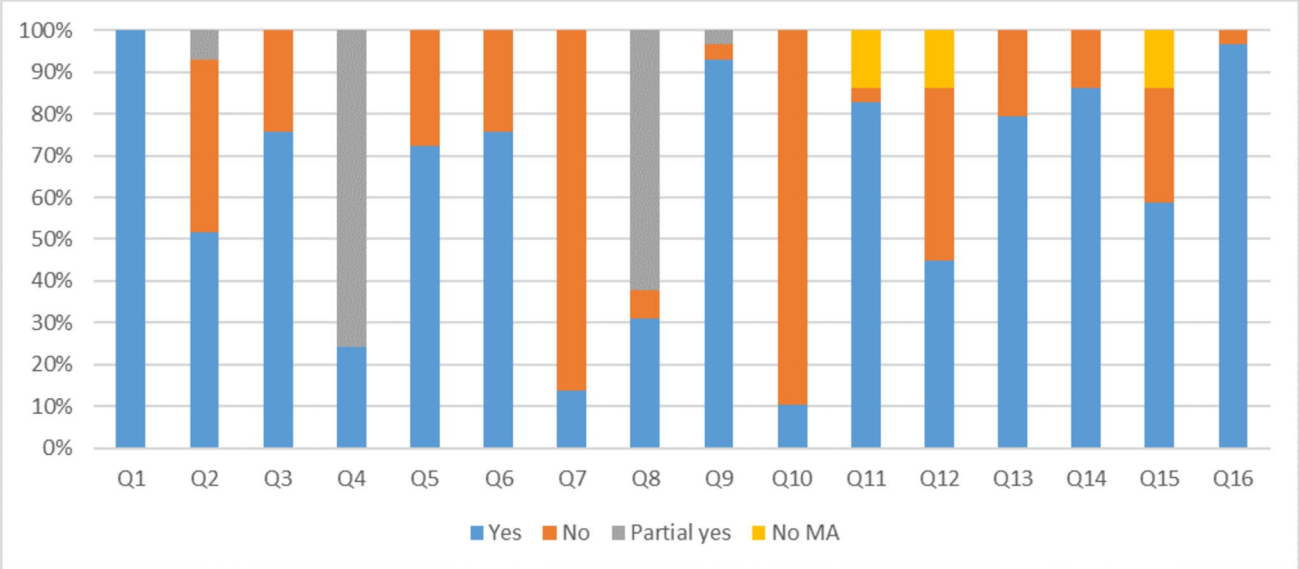


**Fig. 2** Methodological quality appraisal of the included systematic reviews using A MeaSurement Tool to Assess systematic Reviews-2 (AMSTAR-2). Q1 = word research question and inclusion criteria according to PICOS (population, intervention, comparator, outcome, study design), Q2 = establish methods prior to the conduct of the meta-analyses (written protocol), Q3 = explain the choice of study design for inclusion, Q4 = use comprehensive literature search strategy, Q5 = perform study selection in duplicate, Q6 = perform data extraction in duplicate, Q7 = provide a list of excluded studies to justify the exclusion, Q8 = describe the included studies in detail, Q9 = assess the risk of bias, Q10 = reported sources of funding for included studies, Q11 = use appropriate method for statistical combination of results, Q12 = assess the potential impact of risk of bias for included studies, Q13 = account for risk of bias while interpreting/discussing the results, Q14 = explain/discuss any heterogeneity, Q15 = assess publication bias and discuss its impact on the results, Q16 = report potential sources of conflict of interest and describe any funding, "Y" = Yes, "PY" = partial yes, "N" = no, "NA" = not applicable

#### 3.5.4 72 Hours Post-Intervention

With the highest strength of evidence, only kinesiotaping showed a significant effect in reducing pain at 72 h post-intervention (Class III, high RoB in primary studies, $g = 0.94$; [95% CI 1.47, 0.42]) whereas contrast therapy (high RoB in primary studies, $g = 0.38$ [0.74, 0.01]), cooling therapy (moderate RoB in primary studies, $g = 0.78$ [1.23, 0.33]), massage (high RoB in primary studies, $g = 0.79$ [1.42, 0.16]), phototherapy (low RoB in primary studies, $g = 1.13$ [1.89, 0.37]) and vibration therapy (low RoB in primary studies, $g = 1.53$ [2.99, 0.07]) indicated significant low to large effect sizes but a lower strength of evidence (Class IV) (Fig. 6). Electrical stimulation, stretching, cryostimulation, compression therapy, heat therapy and exercise did not significantly reduce DOMS (Class V) (Fig. 7).

#### 3.5.5 96 Hours Post-Intervention

Compression therapy (high RoB in primary studies, $g = 0.38$ [0.71, 0.05]), phototherapy (low RoB in primary studies, $g = 0.84$ [1.40, 0.28]) and contrast therapy (high RoB in primary studies, $g = 1.21$ [2.06, 0.35]) showed a significant effect in reducing pain at 96 h post-intervention (Class IV) whereas cryostimulation, heat or cooling therapy, and massage were not effective in DOMS reduction (Class V) [Fig. 8].

### 3.6 Primary Studies Overlap in SRs

The total number of publications fulfilling the PICO criteria was 303, with 249 RCTs included in the analysis after removing duplicates. The amount of overlap based on the CCA (expressed in %) ranged considerably from slight for compression therapy (0.0%), moderate for cooling therapy (8.8%) and blood flow restriction therapy (11.1%), high for stretching (11%) and electrical stimulation (12.4%) to very high for contrast therapy (15.7%), massage (18.5%), cryostimulation (19.2%), kinesiotaping (22.2%), phototherapy (25%), heat therapy (38.9%) and vibration therapy (33.3%). The tabular and graphical displays of overlap are presented in the ESM.

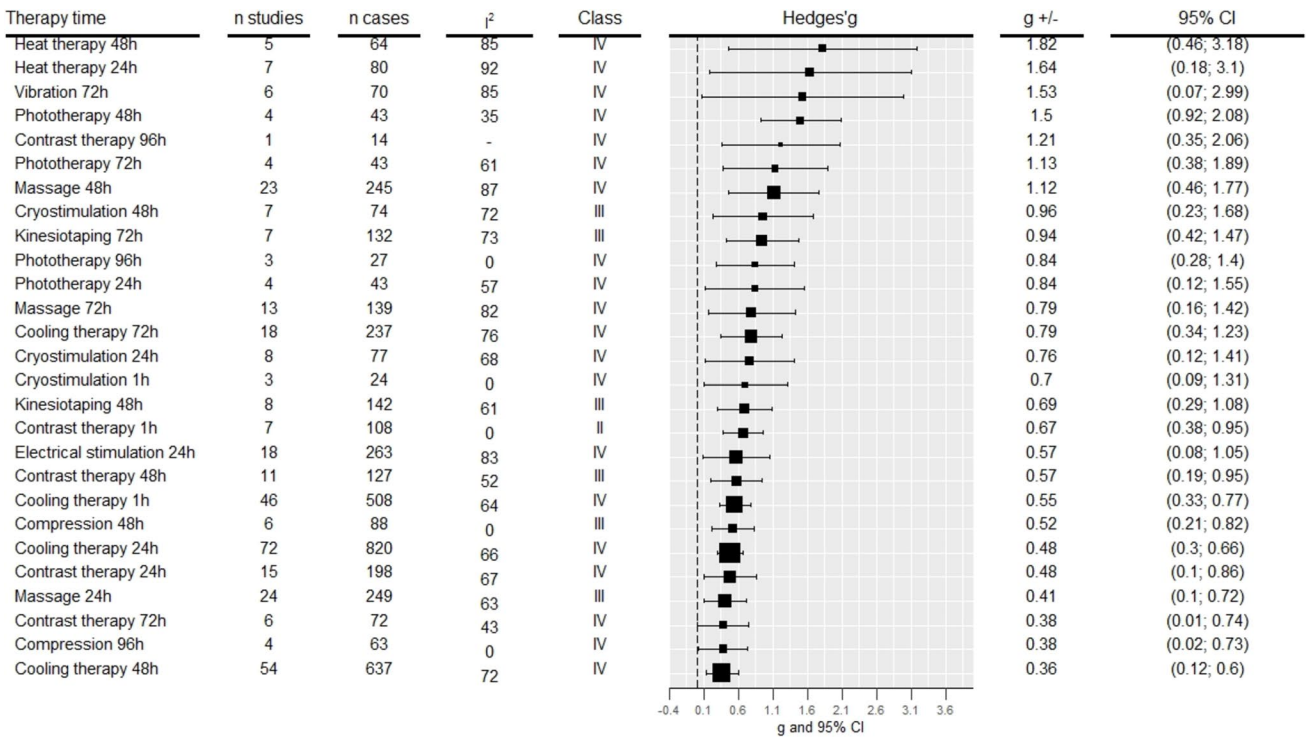


**Fig. 3** Forest plot demonstrating the effectiveness of various therapies on pain intensity indices at specific timepoints post-exercise. Positive Hedges' *g* and 95% confidence interval (CI) values indicate an improvement with therapy interventions over control groups. *h* hours, $I^2$ heterogeneity, *n* studies number of studies in the meta-analysis, *n cases* number of participants in experimental groups in the meta-analysis, Class strength of evidence as described in Sect. 2.5 of the main text. Data presented only for significant main effects (Hedges' *g*), non-significant effects are reported in the ESM

## 4 Discussion

To the best of our knowledge, this is the first UR that summarises and critically evaluates the evidence for the effectiveness of all therapies commonly used by PTs/physical therapists for DOMS.

### 4.1 Meaning of the Findings

We re-analysed findings from almost 900 RCTs and subjected data to a meta-analysis. This created the map of evidence, which is noteworthy for several reasons. First and foremost, it highlights the most effective versus the least effective therapies used, supervised and/or recommended by PTs for treating DOMS. Except for cryo chambers, most of the treatment modalities evaluated here are relatively low cost and can be implemented by PTs. For instance, the evidence was highly suggestive for contrast therapy at three different endpoints, namely immediately after, at 48 and 96 h post-intervention, for compression and kinesiotaping both at 48 and 96 h post-intervention, and for massage and phototherapy at 24 and 96 h post-intervention, respectively. Only one intervention concerning exercise/active recovery at 72 h post-intervention showed excess statistical significance bias under the $p < 0.10$ threshold. This suggests a marginal overrepresentation of statistically significant results compared with what would be expected in the absence of true effects, indicating that publication bias was not present in the majority of the therapies.

We observed inconsistencies in the naming and inclusion of various forms of cold application, as gaseous cryostimulation and cooling therapies are likely to have different physiological effects because of differences in heat capacity and mechanisms of heat exchange [69, 70] Specifically, some studies used the term cryotherapy broadly, [16, 49, 53, 54] while others restricted it to specific interventions such as gaseous cryostimulation, [14, 17, 56] which can induce a rapid short-term decrease in the temperature of the skin and subcutaneous tissues, or cooling methods, [13, 15, 48, 50] which achieve a deeper and longer lasting reduction in tissue temperature. In some cases, there were no clear inclusion criteria or precise definitions for cold-based therapies [52] Additional information with a detailed discussion of the impact of individual therapies on DOMS is included in the ESM.

However, we also identified a considerable amount of heterogeneity in the included SRs and primary trials, which impedes meaningful interpretations and hampers strong

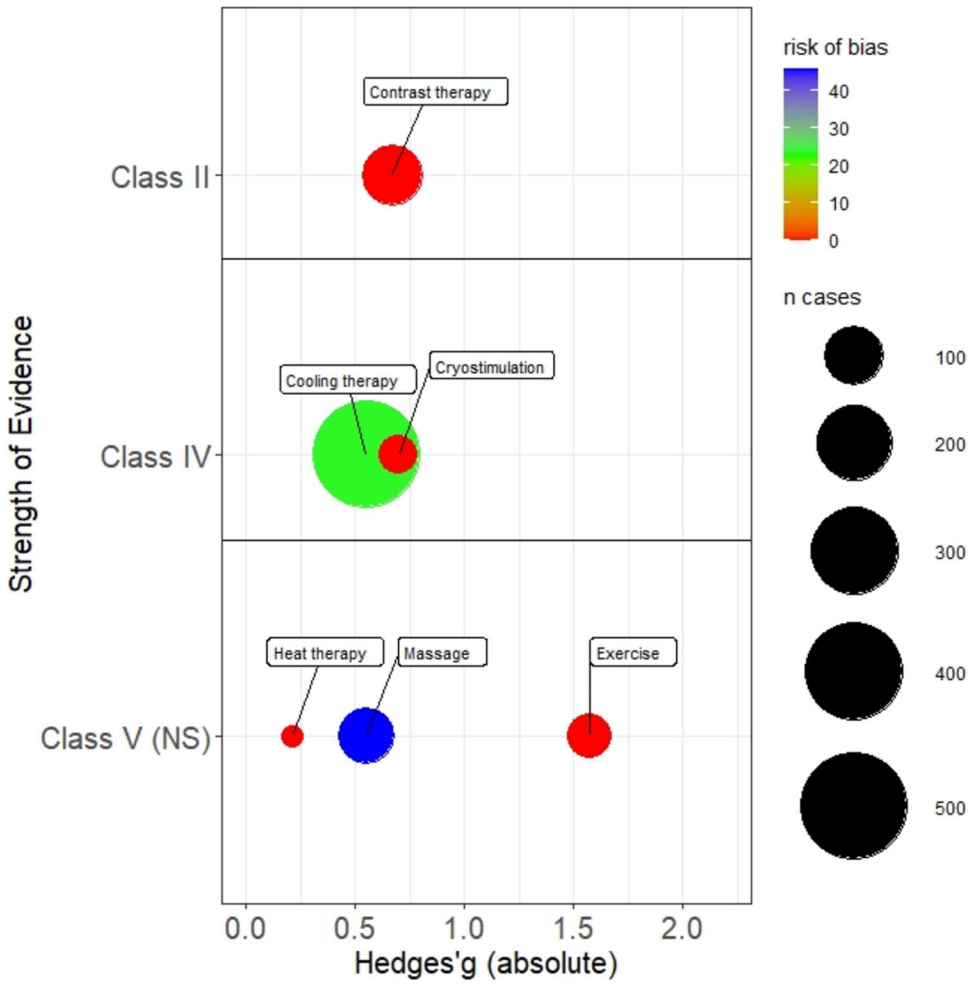


**Fig. 4** Evidence map for the effectiveness of various therapeutic interventions used immediately after exercise for reducing post-exercise muscle soreness. *NS* non-significant main effects (Hedges' *g*)

conclusions. For example, across SRs, I2 would typically be in excess of 75%, which means these findings should not have been pooled in the first instance. There was a wide diversity of populations (and unequal distribution of sex or performance status, i.e. athletes vs non-athletes), duration, intensity or frequency of interventions as well as comparators; heterogeneity of measurement instruments and time to/length of follow-ups; variations of effect sizes (from small to very large) and confounding factors. Presumed mechanisms of action included enhanced lactic acid removal, increased vasoconstriction, release of endorphins or reduced acetylcholine production. Additionally, DOMS is associated with microtrauma to muscle fibres, leading to inflammation and a cascade of repair processes that include increased cytokine production and cellular swelling. The removal of metabolic waste products such as lactate, while important, is only one part of the recovery process. Mechanistically, treatments aimed at preventing or alleviating DOMS may also help to reduce inflammation, limit oxidative stress and modulate immune responses, which contribute to muscle recovery. Furthermore, factors such as diet (determining e.g. antioxidant intake or protein consumption) and sleep patterns can influence these recovery processes by supporting muscle regeneration and reducing inflammation, further impacting the duration and severity of DOMS [2, 5, 8].

We assessed that the methodological quality of the majority of the included SRs was predominantly low or critically low (93% of SRs), which raises questions about the trustworthiness and credibility of those reviews. It is important to note that the AMSTAR-2 assessment includes criteria, such as protocol registration in PROSPERO, that older SRs may not have met. However, the SR by O'Connor and Hurley exhibited significant methodological shortcomings beyond the lack of protocol registration, thus its "critically low'' rating remains justified. It is worth mentioning that guidelines on systematic reviews, such as those by Glasziou et al., were established before the creation of PROSPERO-type registries [71] Furthermore, the quality of the primary studies included in the eligible SRs was predominantly poor and the most prevalent methodological shortcomings included very

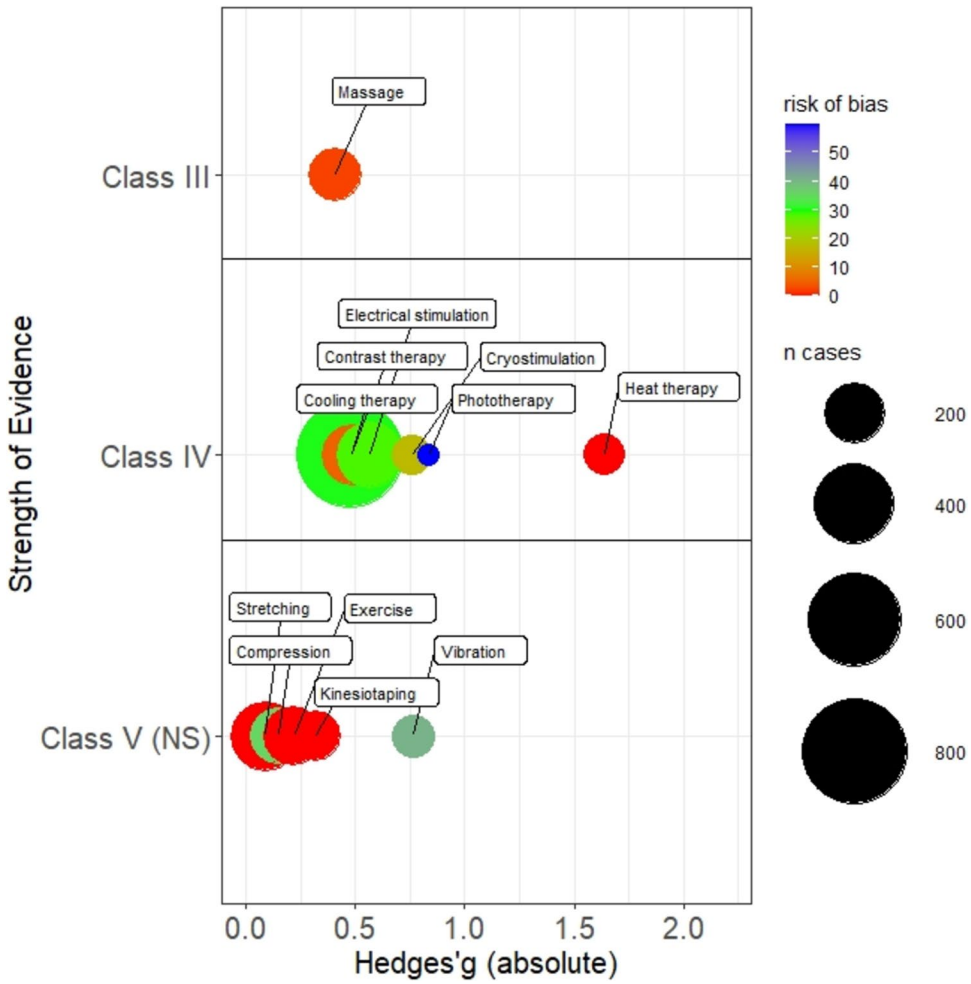


**Fig. 5** Evidence map assessing the effectiveness of different therapeutic interventions utilised post-exercise to alleviate muscle soreness occurring the day after exercise (at 24 h). *NS* non-significant main effects (Hedges' *g*)

small sample sizes, lack of power calculations or blinding (performance bias).

We noticed there was a high interest in the topic as several research groups aimed to synthesise sparse and often overlapping studies. The high overlap between some SRs on heat, vibration, photo or electrical stimulation or massage therapies may lead to several potential misleading conclusions. As search strategies in SRs are rarely reproducible and not always comprehensive, similar SRs might have different included studies [65] When primary studies overlap, i.e. RCTs, only a minority of SR authors have a strategy to deal with this [66] Generally, high interest in the topic can be concluded as several research groups tried to synthesise the apparently sparse and ambiguous primary research.

## 4.2 Comparison with Other Research

Before 1998, mainly narrative/non-SRs on the topic were published. In 1998, Ernst published one of the first SRs evaluating the effectiveness of manual massage in reducing symptoms of DOMS after strenuous exercise [67] He concluded that there was some promising albeit inconclusive evidence suggesting that manual massage is effective in treating DOMS, which is in line with our findings. Of the included SRs, one was a network meta-analysis of 59 trials, which indirectly compared cold and heat therapies for DOMS [17] This SR found heat treatments to be effective within 48 h and cryostimulation beyond 48 h post-intervention; however, it recommended more high-quality trials to draw any firm conclusions.

## 4.3 Limitations

This UR has several strengths including the published protocol, comprehensive searches, data extractions from both secondary and primary studies, statistical pooling of a large number of SRs and creation of the evidence map. Nevertheless, several limitations should be mentioned including publication bias, and an overlap, which is inherent in UR (variable from non-existent to very high). Additionally, we did not include studies published before 1998; however, those have been mentioned in the previous section. We did

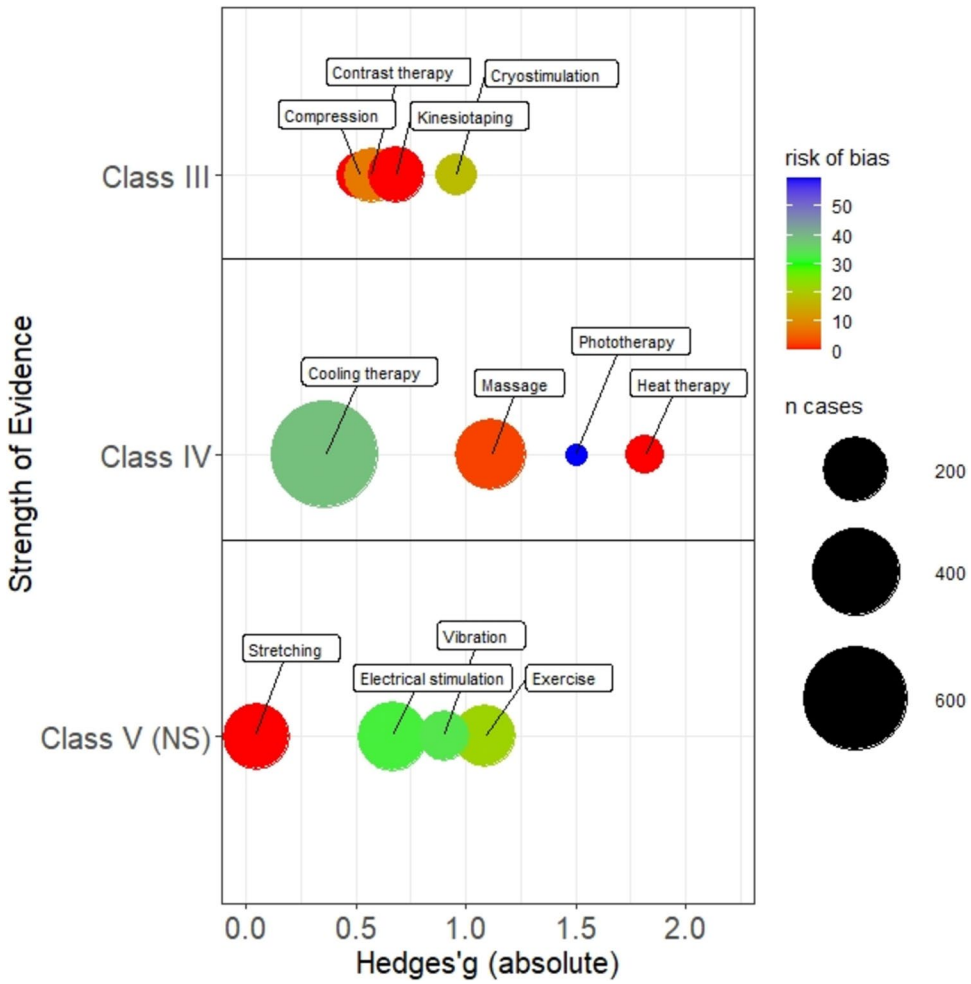


**Fig. 6** Evidence map assessing the effectiveness of different therapeutic interventions utilised post-exercise to alleviate muscle soreness occurring 2 days after exercise (at 48 h). *NS* non-significant main effects (Hedges' *g*)

not include SRs evaluating DOMS in children, adolescents and elderly individuals as well as medically compromised individuals, hence SRs evaluating the effectiveness of any treatments in those groups have yet to be analysed. Our UR is limited to post-exercise interventions for individuals experiencing DOMS, excluding preventive (pre-exercise) interventions from the study's scope. Conducting subgroup analyses was not feasible because of the heterogeneity and poor reporting. We aimed to conduct analyses by types of control groups, i.e. passive versus active, as per the published protocol. However, there is accumulated evidence to suggest that the reporting of superiority, non-inferiority and equivalence trials remains sub-optimal [72–74] As a result, distinguishing between placebo and active controls proved methodologically challenging in our UR.

## 5 Conclusions

There is a large body of evidence from SRs of RCTs evaluating the effectiveness of various physiotherapeutic interventions for DOMS. Nonetheless, the majority of the eligible SRs provide weak and unconvincing evidence, limiting the generalisability of the findings.

### 5.1 Implications for Research

Larger better-quality primary trials could potentially reduce the existing uncertainties. Synergistic effects (if any) of combining different therapies could also be explored. Future research should also explore individual variability in terms of age or fitness level, as well as the role of psychological factors, inflammation or long-term effects of repetitive DOMS episodes.

### 5.2 Implications for Practice

Based on the available evidence, contrast therapy, massage, compression and kinesiotaping appear to be the most promising interventions to alleviate DOMS in healthy adults at follow-ups of up to 96 h post-exercise. These findings are noteworthy for clinicians and trainers who should prioritise

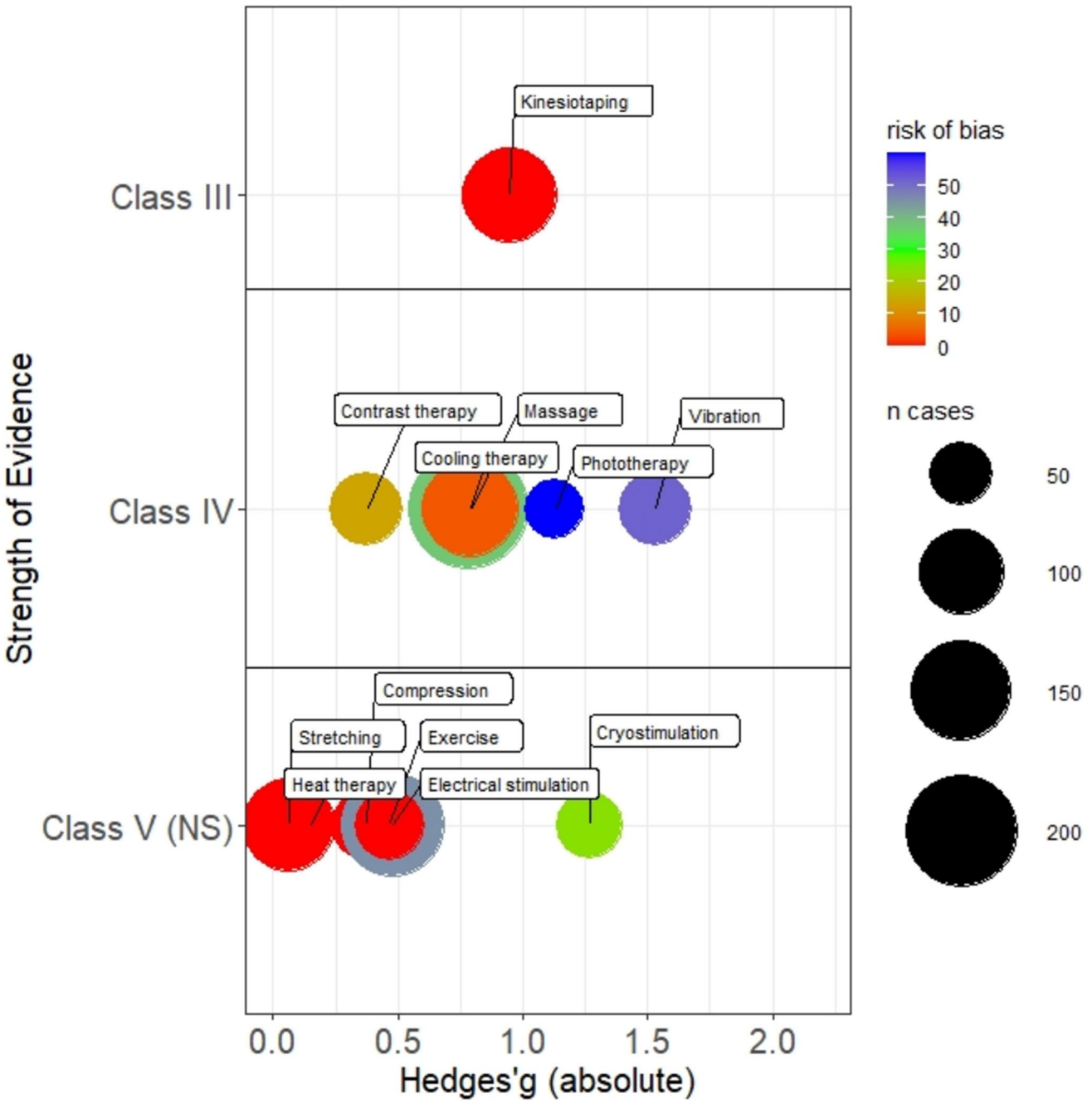


**Fig. 7** Evidence map assessing the effectiveness of different therapeutic interventions utilised post-exercise to alleviate muscle soreness occurring three days after exercise (at 72 h). *NS* non-significant main effects (Hedges' *g*)

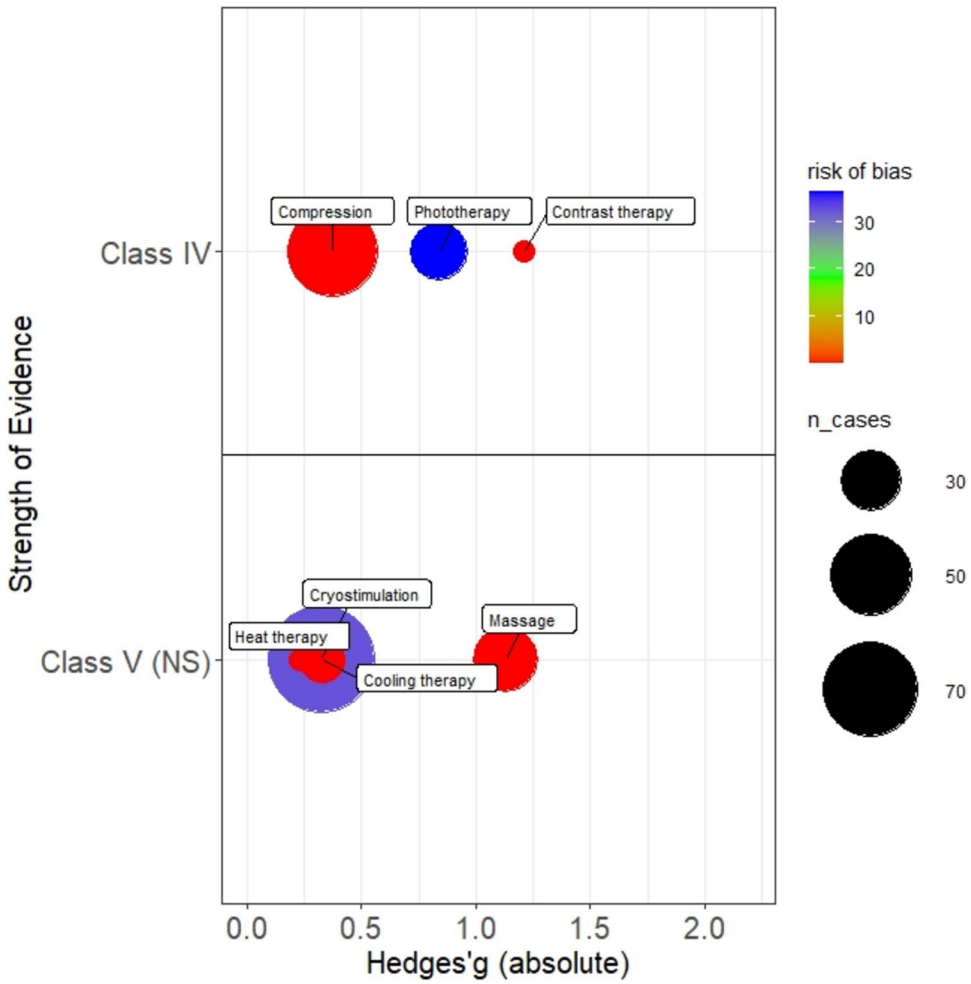


**Fig. 8** Evidence map assessing the effectiveness of different therapeutic interventions utilised post-exercise to alleviate muscle soreness occurring 4 days after exercise (at 96 h). *NS* non-significant main effects (Hedges' *g*)

the use of the most effective therapies to benefit their athletes. Benefits could also be expected in non-athletes.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s40279-025-02187-5.

## Declarations

**Funding** This article was supported by the Ministry of Education and Science, Poland, under the University Research Project, the Jozef Pilsudski University of Physical Education in Warsaw, Poland, grant no. 2/BN/UPB/2023.

**Conflict of Interest** Szczepan Wiecha, Igor Cieśliński, Paweł Wiśniowski, Maciej Cieśliński, Wojciech Pawliczek, Joanna Zając, Paweł Posadzki, Robert Prill and Maciej Płaszewski have no financial or non-financial conflicts of interest that are directly relevant to the content of this article.

**Ethics Approval** Not applicable.

**Consent to Participate** Not applicable.

**Consent for Publication** Not applicable.

**Availability of Data and Material** The datasets analysed in the present study can be obtained on reasonable request from the corresponding author.

**Code Availability** Not applicable.

**Authors' Contributions** Conceptualisation: SW; methodology: SW and MP; pilot screening: SW and RP; screening: RP, PW, MC, WP; data extraction: SW, PW, MC, WP, JZ; random cross-checking: SW, MP; statistics: SW, IC; visualisation: SW, IC; project administration: SW; writing and editing manuscript: SW, MP, PP, RP, JZ; approving final version of manuscript: all authors.

## References

1. Afonso J, Clemente FM, Nakamura FY, Morouço P, Sarmento H, Inman RA, et al. The effectiveness of post-exercise stretching in short-term and delayed recovery of strength, range of motion and delayed onset muscle soreness: a systematic review and meta-analysis of randomized controlled trials. Front Physiol. 2021;12:677581.

2. Cheung K, Hume P, Maxwell L. Delayed onset muscle soreness: treatment strategies and performance factors. Sports Med. 2003;33:145–64.
3. Beaton LJ, Tarnopolsky MA, Phillips SM. Variability in estimating eccentric contraction-induced muscle damage and inflammation in humans. Can J Appl Physiol. 2002;27:516–26.
4. Ulbricht A, Gehlert S, Leciejewski B, Schiffer T, Bloch W, Höhfeld J. Induction and adaptation of chaperone-assisted selective autophagy CASA in response to resistance exercise in human skeletal muscle. Autophagy. 2015;11:538–46.
5. Paulsen G, Mikkelsen UR, Raastad T, Peake JM. Leucocytes, cytokines and satellite cells: what role do they play in muscle damage and regeneration following eccentric exercise? Exerc Immunol Rev. 2012;18:42–97.
6. Sonkodi B, Berkes I, Koltai E. Have we looked in the wrong direction for more than 100 years? Delayed onset muscle soreness is, in fact, neural microdamage rather than muscle damage. Antioxidants (Basel). 2020;9(3):212.
7. Pedersen BK, Ostrowski K, Rohde T, Bruunsgaard H. The cytokine response to strenuous exercise. Can J Physiol Pharmacol. 1998;76(5):505–11.
8. Howatson G, van Someren KA. The prevention and treatment of exercise-induced muscle damage. Sports Med. 2008;38:483–503.
9. Ducas J, Memari S, Houle M, Schwendenmann Y, Abboud J, Yiou É, et al. Impact of lumbar delayed-onset muscle soreness on postural stability in standing postures. Gait Posture. 2024;109:201–7.
10. Schoenfeld BJ. The use of nonsteroidal anti-inflammatory drugs for exercise-induced muscle damage: implications for muscle recovery. Sports Med. 2012;42:1017–28.
11. Ranchordas MK, Rogerson D, Soltani H, Costello JT. Antioxidants for preventing and reducing muscle soreness after exercise. Cochrane Database Syst Rev. 2017;12:CD009789.
12. Clemente-Suárez VJ, Bustamante-Sanchez Á, Mielgo-Ayuso J, Martínez-Guardado I, Martín-Rodríguez A, Tornero-Aguilera JF. Antioxidants and sports performance. Nutrients. 2023;15(10):2371.
13. O'Connor R, Hurley DA. The effectiveness of physiotherapeutic interventions in the management of delayed onset muscle soreness: a systematic review. Phys Ther Rev. 2003;8:177–95.
14. Dupuy O, Douzi W, Theurot D, Bosquet L, Dugué B. An evidence-based approach for choosing post-exercise recovery techniques to reduce markers of muscle damage, soreness, fatigue, and inflammation: a systematic review with meta-analysis. Front Physiol. 2018;9:403.
15. Torres R, Ribeiro F, Alberto Duarte J, Cabri JMH. Evidence of the physiotherapeutic interventions used currently after exercise-induced muscle damage: systematic review and meta-analysis. Phys Ther Sport. 2012;13:101–14.
16. Nahon RL, Silva Lopes JS, Monteiro de Magalhães Neto A. Physical therapy interventions for the treatment of delayed onset muscle soreness (DOMS): systematic review and meta-analysis. Phys Ther Sport. 2021;52:1–12.
17. Wang Y, Lu H, Li S, Zhang Y, Yan F, Huang Y, et al. Effect of cold and heat therapies on pain relief in patients with delayed onset muscle soreness: a network meta-analysis. J Rehabil Med. 2022;54:jrm00258.
18. Batista NP, de Carvalho FA, Machado AF, Micheletti JK, Pastre CM. What parameters influence the effect of cold-water immersion on muscle soreness? An updated systematic review and meta-analysis. Clin J Sport Med. 2023;33:13–25.
19. Guo J, Li L, Gong Y, Zhu R, Xu J, Zou J, et al. Massage alleviates delayed onset muscle soreness after strenuous exercise: a systematic review and meta-analysis. Front Physiol. 2017;8:747.
20. Pieper D, Antoine S-L, Mathes T, Neugebauer EAM, Eikermann M. Systematic review finds overlapping reviews were not mentioned in every other overview. J Clin Epidemiol. 2014;67:368–75.
21. Gates M, Gates A, Pieper D, Fernandes RM, Tricco AC, Moher D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ. 2022;378: e070849.
22. IntHout J, Ioannidis JPA, Borm GF, Goeman JJ. Small studies are more heterogeneous than large ones: a meta-meta-analysis. J Clin Epidemiol. 2015;68:860–9.
23. Miake-Lye IM, Hempel S, Shanman R, Shekelle PG. What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. Syst Rev. 2016;5:28.
24. Biondi-Zoccai G. Umbrella reviews: evidence synthesis with overviews of reviews and meta-epidemiologic studies. Berlin: Springer; 2016. https://doi.org/10.1007/978-3-319-25655-9.
25. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evid Based Ment Health. 2018;21:95–100.
26. Ioannidis JPA. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. CMAJ. 2009;181:488–93.
27. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc. 2015;13:132–40.
28. Wiecha S, Posadzki P, Prill R, Płaszewski M. Physical therapies for delayed onset muscle soreness: a protocol for an umbrella and mapping systematic review with meta-meta-analysis. J Clin Med. 2024;13(7):2006. https://doi.org/10.3390/jcm13072006.
29. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40–6.
30. Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. Lancet Infect Dis. 2010;10(4):226.
31. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358: j4008.
32. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. J Clin Epidemiol. 2019;108:26–33.
33. Duan-Porter W, Coeytaux RR, McDuffie JR, Goode AP, Sharma P, Mennella H, et al. Evidence map of yoga for depression, anxiety, and posttraumatic stress disorder. J Phys Act Health. 2016;13:281–8.
34. Hennessy EA, Johnson BT. Examining overlap of included studies in meta-reviews: guidance for using the corrected covered area index. Res Synth Methods. 2020;11:134–45.
35. Pérez-Bracchiglione J, Meza N, Bangdiwala SI, Niño de Guzmán E, Urrútia G, Bonfill X, et al. Graphical representation of overlap for OVErviews: GROOVE tool. Res Synth Methods. 2022;13:381–8.
36. Gosling CJ, Solanes A, Fusar-Poli P, Radua J. metaumbrella: the first comprehensive suite to perform data analysis in umbrella reviews with stratification of the evidence. BMJ Ment Health. 2023;26(1): e300534.
37. Lin L, Aloe AM. Evaluation of various estimators for standardized mean difference in meta-analysis. Stat Med. 2021;40:403–26.
38. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Behav Stat. 1981;6:107–28.
39. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane handbook for systematic reviews of interventions, 2nd edn. Chichester (UK): John Wiley & Sons, 2019.
40. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 2013.
41. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–58.

42. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34.
43. Stanley TD, Doucouliagos H, Ioannidis JPA, Carter EC. Detecting publication selection bias through excess statistical significance. Res Synth Methods. 2021;12:776–95.
44. Apostolopoulos N, Metsios GS, Flouris AD, Koutedakis Y, Wyon MA. The relevance of stretch intensity and position: a systematic review. Front Psychol. 2015;6:1128.
45. Ma F, Li Y, Yang J, Li X, Zeng N, Martin RL. The effectiveness of low intensity exercise and blood flow restriction without exercise on exercise induced muscle damage: a systematic review. Phys Ther Sport. 2020;46:77–88.
46. Rodrigues S, Forte P, Dewaele E, Branquinho L, Teixeira JE, Ferraz R, et al. Effect of blood flow restriction technique on delayed onset muscle soreness: a systematic review. Medicina. 2022;58.
47. Bieuzen F, Bleakley CM, Costello JT. Contrast water therapy and exercise induced muscle damage: a systematic review and meta-analysis. PLoS ONE. 2013;8: e62356.
48. Bleakley C, McDonough S, Gardner E, Baxter GD, Hopkins JT, Davison GW. Cold-water immersion (cryotherapy) for preventing and treating muscle soreness after exercise. Cochrane Database Syst Rev. 2012;2012:CD008262.
49. Hohenauer E, Taeymans J, Baeyens J-P, Clarys P, Clijsen R. The effect of post-exercise cryotherapy on recovery characteristics: a systematic review and meta-analysis. PLoS ONE. 2015;10: e0139028.
50. Ferlito JV, Ferlito MV, Leal-Junior ECP, Tomazoni SS, De Marchi T. Comparison between cryotherapy and photobiomodulation in muscle recovery: a systematic review and meta-analysis. Lasers Med Sci. 2022;37:1375–88.
51. Xiao F, Kabachkova AV, Jiao L, Zhao H, Kapilevich LV. Effects of cold water immersion after exercise on fatigue recovery and exercise performance: meta analysis. Front Physiol. 2023;14:1006512.
52. Azevedo KP, Bastos JAI, de Sousa Neto IV, Pastre CM, Durigan JLQ. Different cryotherapy modalities demonstrate similar effects on muscle performance, soreness, and damage in healthy individuals and athletes: a systematic review with meta-analysis. J Clin Med Res. 2022;11(15):4441.
53. Nogueira NM, Felappi CJ, Lima CS, Medeiros DM. Effects of local cryotherapy for recovery of delayed onset muscle soreness and strength following exercise-induced muscle damage: systematic review and meta-analysis. Sport Sci Health. 2020;16:1–11.
54. Wang Y, Li S, Zhang Y, Chen Y, Yan F, Han L, et al. Heat and cold therapy reduce pain in patients with delayed onset muscle soreness: a systematic review and meta-analysis of 32 randomized controlled trials. Phys Ther Sport. 2021;48:177–87.
55. Wiśniowski P, Cieśliński M, Jarocka M, Kasiak PS, Makaruk B, Pawliczek W, et al. The effect of pressotherapy on performance and recovery in the management of delayed onset muscle soreness: a systematic review and meta-analysis. J Clin Med Res. 2022;11(8):2077.
56. Costello JT, Baker PRA, Minett GM, Bieuzen F, Stewart IB, Bleakley C. Whole-body cryotherapy (extreme cold air exposure) for preventing and treating muscle soreness after exercise in adults. Cochrane Database Syst Rev. 2015;2015:CD010789.
57. Malone JK, Blake C, Caulfield BM. Neuromuscular electrical stimulation during recovery from exercise: a systematic review. J Strength Cond Res. 2014;28:2478–506.
58. Menezes MA, Menezes DA, Vasconcelos LL, DeSantana JM. Is electrical stimulation effective in preventing or treating delayed-onset muscle aoreness (DOMS) in athletes and untrained adults? A systematic review with meta-analysis. J Pain. 2022;23:2013–35.
59. Lin J, Guo ML, Wang H, Lin C, Xu G, Chen A, et al. Effects of kinesio tape on delayed onset muscle soreness: a systematic review and meta-analysis. Biomed Res Int. 2021;2021:6692828.
60. Nampo FK, Cavalheri V, Ramos S de P, Camargo EA. Effect of low-level phototherapy on delayed onset muscle soreness: a systematic review and meta-analysis. Lasers Med Sci. 2016;31:165–77.
61. Davis HL, Alabed S, Chico TJA. Effect of sports massage on performance and recovery: a systematic review and meta-analysis. BMJ Open Sport Exerc Med. 2020;6: e000614.
62. Herbert RD, de Noronha M, Kamper SJ. Stretching to prevent or reduce muscle soreness after exercise. Cochrane Database Syst Rev. 2011;CD004577.
63. Lu X, Wang Y, Lu J, You Y, Zhang L, Zhu D, et al. Does vibration benefit delayed-onset muscle soreness?: a meta-analysis and systematic review. J Int Med Res. 2019;47:3–18.
64. Yin Y, Wang J, Duan K, Cai H, Sun J. The effect of vibration training on delayed muscle soreness: a meta-analysis. Medicine. 2022;101: e31259.
65. Rethlefsen ML, Brigham TJ, Price C, Moher D, Bouter LM, Kirkham JJ, et al. Systematic review search strategies are poorly reported and not reproducible: a cross-sectional metaresearch study. J Clin Epidemiol. 2024;166: 111229.
66. Gutierrez-Arias R, Pieper D, Lunny C, Torres-Castro R, Aguilera-Eguía R, Oliveros M-J, et al. Only half of the authors of overviews of exercise-related interventions use some strategy to manage overlapping primary studies-a meta-research study. J Clin Epidemiol. 2024;111328.
67. Ernst E. Does post-exercise massage treatment reduce delayed onset muscle soreness? A systematic review. Br J Sports Med. 1998;32:212–4.
68. Jelicic Kadic A, Vucic K, Dosenovic S, Sapunar D, Puljak L. Extracting data from figures with software was faster, with higher interrater reliability than manual extraction. J Clin Epidemiol. 2016;74:119–23.
69. Hohenauer E, Costello JT, Stoop R, Küng UM, Clarys P, Deliens T, et al. Cold-water or partial-body cryotherapy? Comparison of physiological responses and recovery following muscle damage. Scand J Med Sci Sports. 2018;28(3):1252–62.
70. Hohenauer E, Deliens T, Clarys P, Clijsen R. Perfusion of the skin's microcirculation after cold-water immersion (10°C) and partial-body cryotherapy (-135°C). Skin Res Technol. 2019;25(5):677–82.
71. Glasziou P, Irwig L, Bain C, Colditz G. Systematic reviews in health care: a practical guide. Cambridge: Cambridge University Press; 2001.
72. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev. 2012;11(11):MR000030.
73. Piaggio G, Elbourne DR, Pocock SJ, Evans SJ, Altman DG; CONSORT Group. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308(24):2594–604.
74. Kishore K, Mahajan R. Understanding superiority, noninferiority, and equivalence for clinical trials. Indian Dermatol Online J. 2020;11(6):890–4.

## Authors and Affiliations

**Szczepan Wiecha[1] · Igor Cieśliński[2] · Paweł Wiśniowski[2] · Maciej Cieśliński[2] · Wojciech Pawliczek[2] · Paweł Posadzki[3,4] · Robert Prill[5,6] · Joanna Zając[7] · Maciej Płaszewski[2]**

✉ Szczepan Wiecha
szczepan.wiecha@pimmswia.gov.pl

1 Clinical Cardiology Department, National Medical Institute of the Ministry of Interior and Administration, 02-507 Warsaw, Poland

2 Department of Physical Education and Health in Biala Podlaska, Faculty in Biala Podlaska, Jozef Pilsudski University of Physical Education, Warsaw, Poland

3 Faculty of Rehabilitation, University of Physical Culture, Kraków, Poland

4 Kleijnen Systematic Reviews Ltd., York, UK

5 Center of Orthopaedics and Traumatology, University Hospital Brandenburg a.d.H., Brandenburg Medical School Theodor Fontane, Brandenburg an der Havel, Germany

6 Faculty of Health Sciences Brandenburg, Brandenburg Medical School Theodor Fontane, Brandenburg an der Havel, Germany

7 Department of Hygiene and Dietetics, Chair of Epidemiology and Preventive Medicine, Jagiellonian University Medical College, Krakow, Poland