USASP
US ASSOCIATION FOR THE STUDY OF PAIN

PUBLISHED BY
ELSEVIER

The Journal of Pain, Vol 24, No. 12 (December), 2023: pp 2103–2130
Available online at www.jpain.org and www.sciencedirect.com

Featured Article

Trials We Cannot Trust: Investigating Their Impact on Systematic Reviews and Clinical Guidelines in Spinal Pain

Neil O'Connell,[*] R. Andrew Moore,[†] Gavin Stewart,[‡] Emma Fisher,[§] Leslie Hearn,[¶] Christopher Eccleston,[§,‖,**] Michael Wewege,[††,‡‡] and Amanda C. De C Williams[§§]

[*]*Centre for Health and Wellbeing Across the Lifecourse, Department of Health Sciences, Brunel University London, Uxbridge, UK,* [†]*Plymouth, UK,* [‡]*School of Natural and Environmental Sciences, University of Newcastle upon Tyne, Newcastle, UK,* [§]*Centre for Pain Research, University of Bath, Claverton Down, Bath, UK,* [¶]*Cochrane Pain, Palliative and Supportive Care Review Group, Oxford University Hospitals, Oxford, UK,* [‖]*Department of Psychology, University of Helsinki, Finland,* [**]*Department of Clinical and Health Psychology, Ghent University, Belgium, Finland,* [††]*School of Health Sciences, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia,* [‡‡]*Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia,* [§§]*Research Department of Clinical, Educational & Health Psychology, University College London, London, UK*

Abstract: **We previously conducted an exploration of the trustworthiness of a group of clinical trials of cognitive-behavioral therapy and exercise in spinal pain. We identified multiple concerns in 8 trials, judging them untrustworthy. In this study, we systematically explored the impact of these trials ("index trials") on results, conclusions, and recommendations of systematic reviews and clinical practice guidelines (CPGs). We conducted forward citation tracking using Google Scholar and the citationchaser tool, searched the Guidelines International Network library and National Institute of Health and Care Excellence archive to June 2022 to identify systematic reviews and CPGs. We explored how index trials impacted their findings. Where reviews presented meta-analyses, we extracted or conducted sensitivity analyses for the outcomes of pain and disability, to explore how the exclusion of index trials affected effect estimates. We developed and applied an 'Impact Index' to categorize the extent to which index studies impacted their results. We included 32 unique reviews and 10 CPGs. None directly raised concerns regarding the veracity of the trials. Across meta-analyses (55 comparisons), the removal of index trials reduced effect sizes by a median of 58% (Inter Quartlie Range (IQR) 40–74). 85% of comparisons were classified as highly, 3% as moderately, and 11% as minimally impacted. Nine out of 10 reviews conducting narrative synthesis drew positive conclusions regarding the intervention tested. Nine out of 10 CPGs made positive recommendations for the intervention(s) evaluated. This cohort of trials, with concerns regarding trustworthiness, has substantially impacted the results of systematic reviews and guideline recommendations.**

Perspective: ***We found that a group of trials of CBT for spinal pain with concerns relating to their trustworthiness has had substantial impacts on the analyses and conclusions of systematic reviews and clinical practice guidelines. This highlights the need for a greater focus on the trustworthiness of studies in evidence appraisal.***

Pre-registration: ***Our protocol was preregistered on the Open Science Framework: https://osf.io/m92ax/***

Address reprint requests to N.E. O'Connell, PhD, Centre for Health and Wellbeing Across the Lifecourse, Department of Health Sciences, Brunel University London, Uxbridge UB8 3PH, UK. E-mail: neil.oconnell@brunel.ac.uk

Supplementary data accompanying this article are available online at www.jpain.org and www.sciencedirect.com.

1526-5900/$36.00

https://doi.org/10.1016/j.jpain.2023.07.003

Key words: *Clinical trials, Spinal pain, Trustworthiness, Systematic reviews, Clinical practice guidelines*

Randomized Controlled Trials (RCTs) routinely form the basis of evidence included in systematic reviews and meta-analyses of the efficacy or effectiveness of health interventions. They are widely considered the most rigorous form of evidence to guide decisions in Clinical Practice Guidelines (CPGs) and are frequently the only evidence included in systematic reviews of clinical interventions.

RCTs are a human product and so are influenced by biases in human behavior. Evidence-based medicine has numerous tools and methods to assess and manage both quality and bias in research concerned with the conduct and reporting of trials, but there are few methods addressing the important question of trustworthiness of data. Trustworthiness incorporates research integrity and governance, including transparent pre-registration of protocols, appropriate ethical approval and transparent data stewardship, and potential research misconduct.[1] The latter might include fabrication or falsification of research results or plagiarism.[2] If untrustworthy trials are not identified and removed during the development process of reviews and CPGs, then the conclusions and recommendations of those reviews and guidelines are at risk of being incorrect, with a potentially major impact on patient care. This issue is compounded by an academic and publishing system that is generally slow, inefficient and inconsistent in dealing with scientific error, issues of misconduct, and research integrity,[3] and where mistakes are often uncorrected, raising the likelihood of negative impact.[4]

We previously conducted a formal exploration of the trustworthiness of a group of 10 clinical trials of cognitive-behavioral therapy (CBT) and/or physical rehabilitation for persistent spinal pain from a single author team. Three of these trials had been identified as raising concerns during the production of an earlier systematic review[5] and had therefore been excluded. We used several tools designed to examine the plausibility of baseline characteristics and results, as well as searching for preregistration, corrections, or retractions.[1] While we found the index studies had an unremarkable risk of bias profiles, we identified multiple concerns regarding the trustworthiness of 8 of the 10 identified trials (subsequently referred to as index trials)[6–13] (see Supplementary Table 1 for details of those trials, Supplementary Table 2 for a summary of the key domains of trustworthiness that were explored and Supplementary Table 3 for a summary of the concerns raised for each index trial). Key concerns included issues of research governance (lack of study preregistration, no documentary confirmation of relevant ethical approvals, a lack of sharing of data upon request, distributions of baseline variables that appeared unlikely in the context of random allocation, data anomalies (in particular, duplicate or highly similar results data across unique studies), low to no attrition of participants in some studies and implausible results (extremely large effect sizes diverging from the wider literature). On that basis, we recommended that they should not be included in evidence syntheses and clinical guidelines in this area.[1] In this study, we explored how these index trials have affected the conclusions and recommendations of published evidence syntheses and CPGs in spinal pain.

Objectives

To explore the impact of 8 identified clinical trials (index trials) of uncertain trustworthiness on systematic reviews and clinical practice guidelines.

Patient and Public Involvement

Neither patients nor the public were involved in the conception or conduct of this study.

Methods

Our protocol was preregistered on the Open Science Framework: https://osf.io/m92ax/. As this study only involved the evaluation of published information in the public domain, ethical approval was not required.

Inclusion Criteria

We included any systematic review or CPG that cited one or more of the index trials. This included both documents that included the trial(s) in the evidence synthesis and those that identified but excluded the trial(s) from the evidence synthesis.

Search Strategy

We conducted our searches on 22nd and 23rd June 2022 without date restrictions. We used forward citation tracking on the 8 index trials to find systematic reviews and CPGs that included one or more of the trials. For each trial, we searched Google Scholar and used the citationchaser[14] tool (https://estech.shinyapps.io/citationchaser).

We further searched the Guidelines International Network library and the National Institute of Health and Care Excellence (NICE) archive using the terms "back pain" OR "neck pain" to identify any CPGs that may have included the index trials. We planned to search the National Health and Medical Research Council's guidelines portal and the National Guideline Clearinghouse of the Agency for Healthcare Research and Quality but these resources are no longer active.

Two authors independently screened the titles and abstracts of search results, excluding those that were

clearly irrelevant; then they independently screened the full texts of the remaining results, excluding those that were ineligible. Disagreement was resolved by consensus or recourse to a third reviewer.

Citation Summary

From Google Scholar and the citationchaser tool,[14] we explored the number of unique citations received by each trial.

Data Extraction

Data were extracted by one reviewer (NOC) using a standardized form. Table 1 outlines the information included in our extraction.

Data Synthesis

We narratively summarized the affected reviews and CPGs, with details of the affected comparisons. We synthesized narrative results for systematic reviews and CPGs separately. As only one CPG Japanese Orthopedic Association conducted de novo meta-analyses, we included those with our synthesis of meta-analyses from systematic reviews. Where reviews or CPGs made specific practice recommendations on the basis of syntheses that included the index trials, we summarized these.

Where included reviews reported their own sensitivity analyses that excluded the index trials, we extracted those. Where reviews did not report any such analysis and where adequate data were available in reports or from authors on request, we conducted sensitivity meta-analyses for the outcomes of pain and disability at all available follow-up time points, to explore how the exclusion of the index trials would affect published effect estimates. We used Revman5[15] to replicate the published analyses with those trials excluded. We replicated analyses using information presented in reviews regarding the meta-analysis model and compared the results of our analyses with the published analyses for any unexpected divergence that might result from different model parameters. We calculated the proportion of participants that index trials contributed to meta-analyses, the absolute and percentage change in effect sizes for each analysis, and the change in effect size that resulted from the exclusion of the trial(s). To allow comparison of effect sizes across different scales we converted all to the same direction, as they all represented symptomatic improvement. We explored whether the exclusion of the trials changed heterogeneity in the meta-analysis by examining the absolute change in the I^2 statistic (while not a direct measure of heterogeneity, the I^2 statistic measures the percentage of variability in effect estimates due to between-trial heterogeneity),[16] it also has the benefit of being commonly reported in meta-analyses. For all analyses, we used the standardized mean difference (Hedge's g) to maximize comparability. It should be noted that we

Table 1. Details of Data Extracted From Reviews and CPGs

Characteristics of the systematic review/CPG	The overall aims/objectives of the review or guideline The country of origin of the author team/CPG The organization developing the CPG Any declared conflicts of interest of review authors Any declared external funding. Date of publication and journal/policy body
Details of Index trial(s) inclusion in review CPG	The index trials cited in the review/CPG and how they were cited (eg, in background, discussion, or as part of the methods and results of the review/CPG) The index trials excluded from the evidence synthesis in the review/CPG and any reported reasons for exclusion. The index trials included in the evidence synthesis in the review/CPG
Details of Index trials inclusion in analysis/ synthesis.	From the index trials included in the evidence synthesis, we extracted: In which comparison(s) For which outcome(s) Details of affected comparison(s) Whether identified trial(s) were identified/recognized as divergent/ outliers Details of any reported narrative addressing that issue Meta-analysis (pairwise or network) Model details, number of trials, participants, weight given to trials, heterogeneity estimates, pooled effect size/precision estimates. Any exploration (details and results) of heterogeneity (subgroup analysis, sensitivity analysis, meta-regression Whether identified trial(s) were identified/recognized as divergent/ outliers. Details of any evaluation of the certainty of results in affected comparisons (eg, using GRADE) Narrative synthesis. Number of trials, participants. Overall results. Details of any evaluation of the certainty of results in affected comparisons (eg, using GRADE) Conclusions/recommendations Overall conclusions of the evidence synthesis relating to affected comparisons (ie, that include identified trial(s)) Any specific clinical recommendations made in reviews or CPGs that include the index trials

Abbreviation: CPG, clinical practice guidelines.

conducted these analyses using the data reported in the included reviews, and did not go back to the original trials. All re-analyses were checked for errorsby a second researcher.

While not planned in our original protocol, we developed an index to categorize the extent to which index studies had impacted upon the results of identified meta-analyses. This was developed de novo through discussion and consensus among the team and we labeled this the "Impact Index". We first considered and agreed as a team the key domains in which studies may have an impact and then agreed, through discussion and consensus what we considered to be defensible thresholds for judging the impact of index trials in each domain. These were then applied consistently across all included studies. We have not conducted a formal validation of this index. The Impact Index classifies the impact that index studies have had on a meta-analysis across 4 domains, as follows:

Scale of Contribution

What is the proportional volume of data or weight that index trials contribute to the meta-analysis/pooled effect? We quantified this as the proportion of total participants in the meta-analysis that is contributed by index trials, though weight (%) could also be used.

Impact on the Pooled Effect

What is the impact of index trials on the pooled estimate of treatment effect? This was quantified by calculating the absolute change in the treatment effect (Standardized Mean Difference) observed when index trials were excluded from the analysis.

Impact on Precision

What is the impact of index trials on the precision of the effect estimate? This was quantified by calculating the % change in the width of the 95% confidence intervals of the pooled effects when the index trials were excluded from the analysis or where the exclusion of index trials altered the significance status of the effect at the $P < .05$ level. While we recognize that using P values as a marker of meaning is unsatisfactory, it is almost universally used in primary or secondary reporting of meta-analytic outcomes.

Impact on Inconsistency

What is the impact of index trials on inconsistency in the meta-analysis? This was quantified by calculating the absolute change in the I^2 statistics value when index trials were excluded from the analysis or where the exclusion of index trials changed the direction of the effect of the point estimate.

For each domain, judgments were made on whether the impact was substantial, moderate, or low. The decision thresholds for each domain are presented in Table 2. The results for each domain were then combined using a decision rule to produce an overall judgment of the level

Table 2. Decision Thresholds for Each Domain of the Impact Index

	Scale of contribution	*Effect magnitude SMD*	*Precision*		*Inconsistency*	
	Overall proportion of participants from index trials in meta-analysis	*Absolute change in ES*	*Relationship with the null*	*% change in CI margin*	*Absolute change in I^2 statistic*	
Low impact	<15%	<0.2		< 20%	<20%	
Moderate impact	15–39%	0.2–0.49		20–49%	20–49%	
High impact	>40%	≥0.5	Changes "significance status": Changes CI to include no effect or to exclude no effect at the $P < .05$ level	≥ 50%	≥50% or changes direction of point estimate	

Abbreviation: SMD, Standardized Mean Difference.

of impact. The decision rules for overall judgments were as follows:

Highly Impacted

- Substantial impact on the scale of contribution ± any effect on the magnitude of effect, precision, or inconsistency
- Substantial impact on the magnitude of effect ± any effect on the scale of contribution, precision, or inconsistency
- Moderate impact on the scale of contribution ± moderate or substantial effect on the magnitude of effect, precision, or inconsistency
- Moderate impact on the magnitude of effect ± moderate or substantial effect on the scale of contribution, precision, or inconsistency

Moderately Impacted

- Moderate impact on the scale of contribution with minimal impact on the magnitude of effect, precision, and inconsistency
- Moderate impact the magnitude of effect with minimal impact on the scale of contribution, precision, and inconsistency
- Minimal impact on scale of contribution or magnitude of effect with moderate or substantial impact on precision and/ or inconsistency

Minimally Impacted

Minimal effect on the scale of contribution, effect magnitude, precision, or inconsistency.

Results

Fig. 1 presents the results of the searches and screening process. In summary, citation tracking identified 729 records, and guideline database searches identified 147 records. After removal of duplicates and screening of titles and abstracts, we reviewed 81 records of which we excluded 39 (for reasons see Fig. 1). We finally included 42 records (32 unique systematic reviews[5,17–47] and 10 CPGs[48–57]).

For the individual index trials, the number of unique citations ranged from 4 to 143. The number of identified systematic reviews that an individual trial was included in ranged from 0 to 16, and the number of CPGs that an individual trial was included in ranged from 0 to 4. The Monticone et al (2013) study[7] was the most cited, included in most systematic reviews, and, along with the Monticone et al (2016) study,[9] was included in most CPGs. Fig. 2 summarizes these data.

Impact on Systematic Reviews

Table 3 presents the characteristics of the included systematic reviews. These were published between 2015 and 2022 and were conducted by author teams from a variety of countries. Most focused on the effectiveness of psychological, multidisciplinary, or multicomponent

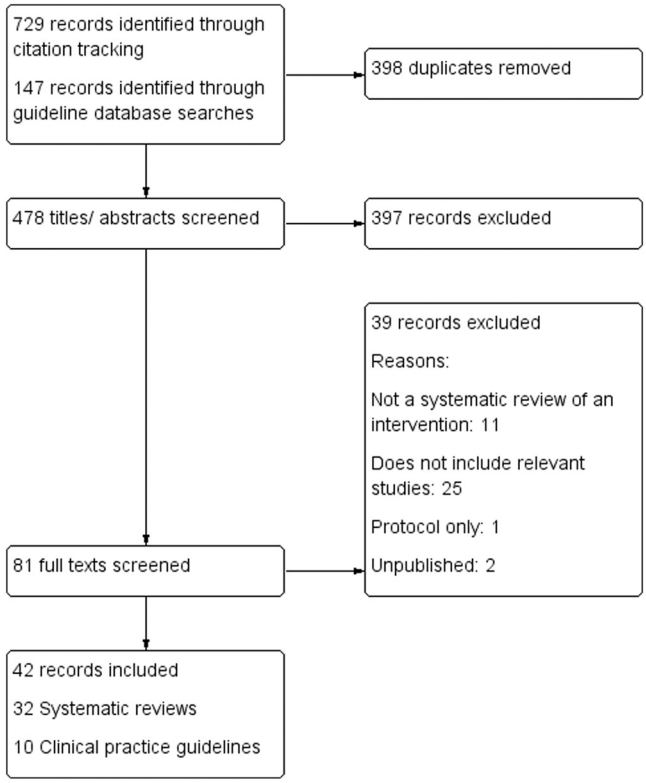

Figure 1. Flow diagram of the search screening process.

interventions for persistent pain[5,18–23,25,27–32,35,38–41,43–46] or postsurgical pain.[23,33,34,36,37] Two[17,47] focused on the effectiveness of stabilization exercises for back pain in people with scoliosis) and one on the effectiveness of exercise interventions for chronic low back pain.[26] Ten reviews reported using GRADE to evaluate the certainty of evidence.[19,20,23,24,26,29,31,32,40,42]

Impact on Meta-Analyses

Of these reviews and 1 CPG,[53] 21 conducted pairwise meta-analyses and 2 undertook network meta-analyses. We were able to extract the results of sensitivity analyses where index trials were excluded from 4 reviews,[5,19,20,24] of which one[20] provided additional data on request. One review with an included NMA[27] reported the results of such a sensitivity analysis conducted after a rapid response to their review by the authors of this paper, raising concerns about the inclusion of identified/index trials, and we used those reported results. In one included review,[36] data in the reported analysis did not match the reported outcome data from the included studies which indicated that it was erroneous. As that review reported the necessary data, we reconducted the meta-analysis with the relevant trials included and then excluded. For all other analyses, we encountered no unexpected divergence between the results of our analyses with the index trials included and those reported in the included reviews.

In total, we included 55 sensitivity analyses of the impact of excluding the index trials from meta-analyses, of which 16 were conducted by the authors of the

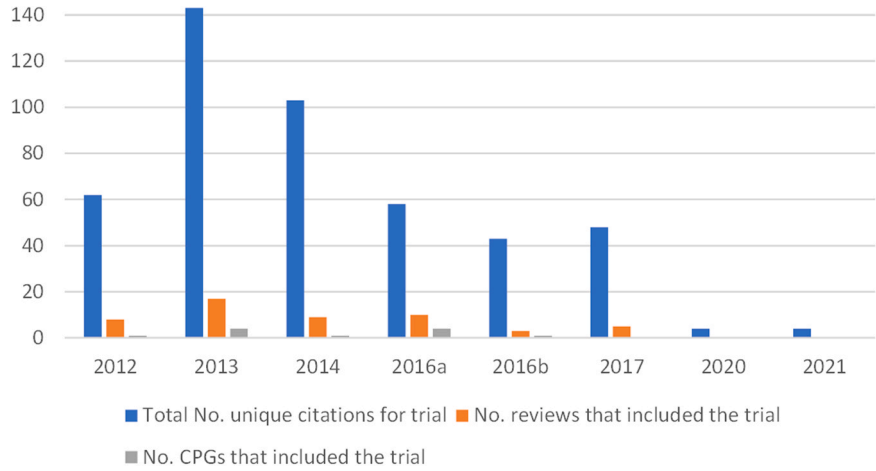

Figure 2. The number of unique citations, and appearances in systematic reviews and CPGs for each index trial. *Footnotes: Each date represents a unique index trial, identified by its year of publication.*

reviews,[5,19,20,24] and 39 were conducted by us. Table 4 summarizes these results.

Across all meta-analyses for pain and disability at all time points (55 comparisons), the removal of the index trials reduced effect sizes by a median of 58% (Inter Quartlie Range (IQR) 40–74). This reflected a reduction in effect size for all comparisons. The median absolute reduction in effect size (Standardized Mean Difference) was –0.35 (IQR –0.51 to –0.21). We saw the same pattern of reduced effect sizes when we grouped analysis by outcome (pain or disability) or clinical population (chronic pain or postsurgical pain). Fig. 3 illustrates the absolute reduction in effect size observed when index trials were removed from each analysis. Even a proportionally small contribution from index trials commonly resulted in moderate to large differences in the pooled effect size in a substantial number of cases.

Removal of index trials also led to a reduction of heterogeneity (Table 4); the I^2 statistic across all meta-analyses for pain and disability at all time points (44 comparisons) reduced by a median of 27% (IQR 10–66). A similar pattern was observed whether we grouped analysis by outcome (pain or disability) or by clinical population (chronic pain or postsurgical pain). In terms of impact on precision, the removal of index trials led to a reduction in the width of the 95% confidence interval in 48 comparisons, no change in 2, and an increase in the width in 5. The exclusion of index trials reduced the width of the confidence intervals by 48% (27–65). This increase in precision after exclusion reflects the divergent nature of the effect sizes reported in these trials. These data are presented in Table 4. Twelve out of 40 statistically significant effects (at the $P < .05$ threshold) became nonsignificant after the removal of studies of interest.

Application of the Impact Index resulted in 47 of the 55 comparisons (85%) being classified as highly impacted by the inclusion of index trials, 2 (4%) as moderately impacted and six (11%) as minimally impacted. Sensitivity analyses using study weight in place of the proportion of participants contributing to judge the "scale of contribution" domain did not alter any overall impact judgments. Judgments by domain for each included comparison are presented in Supplementary Table 4. As noted above, 4 reviews conducted and reported their own analyses examining the impact of excluding the index studies from their analysis[19,20,24] or by excluding index studies from the primary analysis.[5] On that basis, we considered these reviews to have clearly addressed the potential impact of the index trials.

In 15 of these reviews, authors commented on the fact that studies of interest were outliers in their sample, had very large effects, and/or introduced heterogeneity to the analyses.[5,19–21,23,24,29,33,34,39,40,42,44–46] Of these, the authors of 5 reviews[19,23,24,44,45] speculated that the dose, intensity, and/or aspects of the content of the intervention in those trials might explain the observed divergence, while the other reviews either did not offer an explanation or stated that the heterogeneity was unexplained.

Impact on Narrative Syntheses

Ten reviews[17,18,25,28,30,31,34,41,43,44] did not conduct meta-analyses but synthesized the evidence narratively. Nine of these reported the total number of trials (range 1–61), and participants (range 80–7201) included in the review. The proportion of participants contributed by index trials ranged from 4.4% to 100% (median 9.3%). Nine of those reviews[17,18,25,28,30,31,34,43,44] drew broadly positive conclusions regarding the effectiveness of the intervention under scrutiny, underpinned in part by evidence from index trials. Only 2 specifically commented on the heterogeneity in the results of included trials or specifically referred to the divergent results of the index trials. Of these, 1 review[28] referred to the Monticone et al[7] trial as an example of the potential benefits of group-based therapy and another commented that studies with a higher number of treatment sessions (which included Monticone et al[7] and 2016a[9]) found larger effects on pain intensity than those with fewer sessions. No review raised specific concerns regarding the veracity of the data from index trials.

Table 3. Characteristics of Reviews That Included the Index Trials

Authors, primary	*Country of origin*	*Review primary objectives*	*Search date*	*Index trials included*	*Comparisons affected*	*meta-analysis Y/N*	*narrative synthesis Y/N*	*GRADE Y/N*	*Overall review conclusions*
Alanazi et al 2018[17]	Canada, Saudi Arabia	".to evaluate the effect of the stabilization exercise on back pain in adults with scoliosis"	To March 9, 2017	10	Active self-correction, task-oriented exercises, and cognitive-behavioral therapy versus "general physiotherapy"…	N	Y	N	"Stabilization exercise, as reported in the included study, is shown to be effective in reducing back pain, disability and improving quality of life in adults with idiopathic scoliosis. However, this review highlights the paucity of literature examining the effect of exercise on back pain in adults with scoliosis and strongly suggests that further experimental research is needed aiming to ensure proper blinding as this was a common weakness."
Barbari et al 2020[18]	Italy	"to elaborate the state of the art of scientific literature on the effectiveness of interventions that included communicative and educative strategies on 3 main outcomes: 1) patient's LBP awareness/knowledge, 2) maladaptive behavior modification and 3) Compliance with exercise"	Between September and February 2018	7,9	CBT + physiotherapy versus physiotherapy alone	N	Y	N	"Evidence coming from high-quality RCTs sustains multimodal interventions, pain science education and graded exposure—as well as combinations of CBT, pain science education and graded exposure—on behavior modification in the mid-and long-term."
Bogaert et al 2022[19]	Belgium	"to assess and compare the effectiveness of unimodal and multimodal rehabilitation strategies on disability, pain, and pain-related fear in patients undergoing lumbar fusion surgery for degenerative conditions and (adult) isthmic spondylolisthesis. The secondary aim was to assess the effectiveness of return-to-work."	To April 2021	8	Multimodal rehabilitation intervention versus exercise therapy	Y	N	Y	"The results of this systematic review with meta-analysis encourage exercise for all patients undergoing lumbar fusion given the positive impact on disability and pain up to 6 months postoperative. Embedding exercise in a multimodal rehabilitation context is suggested given the additional positive effect on disability and pain-related fear, compared to exercise alone. It remains uncertain if these beneficial effects of exercise and multimodal rehabilitation persist in the long term."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Casey et al 2020[20]	Ireland	Systematic review examining the evidence for MBR compared with active physical interventions in chronic noncancer pain conditions	inception to November 2018	7,9,11	MBR versus active physical treatment	Y	N	N	"Low-quality evidence to suggest that MBR is more effective than active physical interventions for reducing pain intensity and disability in people with chronic pain, in both the short term and long-term. As the benefits appear to be modest, other factors such as resource allocation should be considered before offering MBR programs to people with chronic pain."
Cheng et al 2019[21]	UK, Hong Kong	"to examine the effects of physical exercise cum cognitive-behavioral therapy (CBT) on alleviating pain intensity, functional disabilities, and mood/mental symptoms in those suffering with chronic musculoskeletal pain."	To December 31, 2018	6,7,9,11	CBT +exercise versus exercise alone	Y	N	N	"The value of adding CBT to exercise interventions is questionable, as consistent benefits were not seen. The clinical implications and directions for future research are discussed."
Fadli et al 2021[22]	Indonesia	"to summarize the existing literature data regarding the effectiveness of biopsychosocial interventions with Cognitive Behavioral Therapy (CBT) and Exercise Therapy Program (ETP) in CLBP."	To 2020	7,9	CBT + exercise versus exercise alone	Y	N	N	"Biopsychosocial interventions for CBT and ETP are effective in reducing the pain and disability index of CLBP. The effectiveness of biopsychosocial interventions with CBT and ETP methods is due to the patient's efficacy in completing the rehabilitation program."
Greenwood et al 2016[23]	UK	"to conduct a systematic review and meta-analysis of current evidence evaluating the effectiveness of rehabilitation following lumbar fusion surgery (LFS)."	To October 13, 20, 2014	8	Exercise +CBT versus exercise alone	Y	N	N	"A small number of low-quality studies suggest that ''complex rehabilitation'' reduces short and long-term disability and fear avoidance behavior following LFS. More, high-quality research is required to confirm the effectiveness of ''complex rehabilitation'' programs."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Guerrero et al 2018[24]	Australia	"we examined the effectiveness of physiotherapist delivered psychological interventions combined with physiotherapy on pain, disability and psychological outcomes for patients with musculoskeletal pain conditions"	To May 2016	6	Combined psychological and physiotherapy versus physio/ usual care alone	Y	N	Y	"The results indicate that psychological interventions delivered by physiotherapist show promise to improve health outcomes, particularly psychological outcomes, in musculoskeletal pain conditions."
Hajihasani et al 2019[25]	Iran	"To investigate the effect of adding the cognitive-behavioral treatment (CBT) component to routine physical therapy (PT) on pain and depression reduction, improvement in quality of life, and enhanced function in patients with chronic low back pain (CLBP)."	To January 2018	7	CBT + physio versus physio alone	N	Y	N	"although some patient populations benefited from receiving CBT in addition to routine PT in their CLBP rehabilitation process in terms of reducing pain and disability and enhancing functional capacity and quality of life, there were others for whom CBT did not seem advantageous. None of the investigations found that the addition of CBT assisted in reducing depression symptoms."
Hayden et al 2021[26]	Canada	"The primary objective of this systematic review is to assess the impact of exercise treatment on pain and functional limitations in adults with chronic non-specific low back pain compared to no treatment, usual care, placebo and other conservative treatments."	To April 27, 2018	9	?	Y	N	Y	"We found moderate-certainty evidence that exercise is probably effective for treatment of chronic low back pain compared to no treatment, usual care or placebo for pain. The observed treatment effect for the exercise compared to no treatment, usual care or placebo comparisons is small for functional limitations, not meeting our threshold for minimal clinically important difference. We also found exercise to have improved pain (low-certainty evidence) and functional limitations outcomes (moderate-certainty evidence) compared to other conservative treatments; however, these effects were small and not clinically

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
									important when considering all comparisons together."
Ho et al 2022[27]	Australia	"To determine the comparative effectiveness and safety of psychological interventions for chronic low back pain."	To January 2021	7,9	NMA in node "cognitive behavioral therapy delivered with physiotherapy care"	Y NMA	N	N	"For people with chronic, non-specific low back pain, psychological interventions are most effective when delivered in conjunction with physiotherapy care (mainly structured exercise). Pain education programmes (low to moderate quality evidence) and behavioral therapy (low to high quality evidence) result in the most sustainable effects of treatment; however, uncertainty remains as to their long term effectiveness."
Ince et al 2020[28]	Turkey	"to systematically review the studies investigating the effectiveness of all forms of CBT for the treatment of individuals with chronic pain."	To February 2014	6,7	Not clearly reported	N	Y	N	"In conclusion, besides demonstrated effectiveness of cognitive behavioral therapies, there is no clear explanation for which specific treatment components are responsible for the improvement in which specific complaints."
Kamper et al 2015[29]	Australia	"To assess the long-term effects of multidisciplinary biopsychosocial rehabilitation for patients with chronic low back pain."	To February 2014	7	Multidisciplinary biopsychosocial rehab versus physical treatment	Y	N	Y	"Patients with chronic LBP receiving MBR are likely to experience less pain and disability than those receiving usual care or a physical treatment. MBR also has a positive influence on work status compared to physical treatment. Effects are of a modest magnitude and should be balanced against the time and resource requirements of MBR programs. people with indicators of significant psychosocial impact are referred to MBR."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Knoerl et al 2015[30]	USA	"to determine (a) which CBT doses, delivery methods, strategies, and follow-up periods have been explored in recent intervention studies of individuals with chronic pain"	Between 2005 and 2015	6,7	CBT	N	Y	N	"The results of this review demonstrated that CBT was effective for pain intensity in 43% of the trials and was an effective treatment for many pain-related variables recommended by IMMPACT such as physical functioning, anxiety, depression, and quality of life."
Martinez-Calderon et al 2020[31]	Spain	"To systematically review and critically appraise the effectiveness of conservative and surgical interventions to reduce fear in studies of people with chronic low back pain, based on the analysis of randomized controlled trials for which fear was a primary or secondary outcome."	To May 2019	7–9	Multidisciplinary interventions	N	Y	Y	"Multidisciplinary and psychological interventions as well as exercise reduced kinesiophobia. Fear-avoidance beliefs were reduced by the aforementioned interventions, manual therapy, and electrotherapy. A multidisciplinary intervention reduced the fear of falling. There was moderate evidence of multidisciplinary interventions and exercise to reduce kinesiophobia."
Monticone et al 2015[32]	Italy	"To assess the effects of CBT among individuals with subacute and chronic NP. Specifically, the following comparisons were investigated: 1) cognitive-behavioral therapy versus placebo, no treatment, or waiting list controls; 2) cognitive-behavioral therapy versus other types of interventions; 3) cognitive-behavioral therapy in addition to another intervention (eg, physiotherapy) versus the other intervention alone."	To November 2014	6	CBT in addition to another intervention versus the other intervention alone. CBT in addition to another intervention versus the other intervention alone	Y	N	Y	"When comparing both CBT to other types of interventions and CBT in addition to another intervention to the other intervention alone, no differences were found."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Nadinda et al 2022[33]	Belgium	"to investigate the efficacy of perioperative psychological interventions in reducing (sub) acute postsurgical pain and CPSP and disability in adults."	To March 2020	8	Psychological interventions versus control (+moderator analyses)	Y	N	N	"Psychological interventions can be effective in reducing (sub) acute postsurgical pain and CPSP and disability. These results underscore the possible benefits of integrating psychological services into multidisciplinary acute and transitional pain teams.implementing psychological interventions in the perioperative period may have the potential to reduce the humanitarian and economic burden of CPSP."
Nicholls et al 2018[34]	Canada	"To identify randomized controlled trials evaluating the efficacy of these psychotherapy approaches on pain-related surgical outcomes."	To 2017	8	Effects of CBT on pain-related outcomes	N	Y	N	"This systematic review provides preliminary evidence that CBT-based psychological interventions reduce PSP intensity and disability. Future research should further clarify the efficacy and optimal delivery of CBT and newer psychological approaches to PSP."
O'Keeffe et al 2016[35]	Ireland	"to assess the comparative effectiveness of physical, behavioral and/or psychologically informed, and combined interventions on pain and disability in patients with NSCSP."	To January 2013	6	Effect of physical versus combined interventions on pain and disability	Y	N	N	"No clinically significant differences were found for pain and disability between physical, behavioral/ psychologically informed, and combined interventions for NSCSP."
Özden et al 2022[36]	Turkey	"to systematically review the effect of exercise interventions and conduct its meta-analysis in patients after LFS."	To October -December 2021	8	cognitive behavioral training in addition to exercise versus exercise	Y	N	N	"The results of the present systematic review and meta-analysis reported that cognitive therapy applied in addition to exercise could provide more effective results in physical and psychological parameters."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Parrish et al 2021[37]	USA	"to investigate the influence of CBT on patient reported outcomes among lumbar spine surgery patients."	To December 2019	8	The effect of CBT on back pain, leg pain, disability versus usual care	Y	N	N	"Compared to usual care or alternative therapy control arms, CBT delivered the most improvement with overall quality of life and psychological outcomes. Among appropriately selected patients, CBT could improve perioperative disability, pain, quality of life, and psychological health following lumbar spine surgery."
Petrucci et al 2022[38]	Italy	"to identify and to describe the most common psychological approaches used to treat patients who suffer from CLBP"	Not reported	7	CBT versus control	Y	N	N	"CBT and MBSR have proven their significant effectiveness to improve pain intensity and quality of life compared to controls. These approaches also demonstrated their efficacy in reducing disability and fear-avoidance, but without significant results. Our findings suggest that CBT and MBSR modify pain-related outcomes and that they could be implemented in clinical practice."
Richmond et al 2015[39]	UK	"To assess whether cognitive behavioral (CB) approaches improve disability, pain, quality of life and/or work disability for patients with low back pain (LBP) of any duration and of any age."	To November 2014	7	CBT versus guideline-based active treatment	Y	Y	N	"CB interventions yield long-term improvements in pain, disability and quality of life in comparison to no treatment and other guideline-based active treatments for patients with LBP of any duration and of any age."
Schütze et al 2018[40]	Australia	"to examine all treatment-related changes in PC regardless of whether catastrophizing was specifically targeted as a primary outcome."	To November 2016	8–11	Multimodeal treatment versus active control	Y	N	Y	"The best evidence (moderate-high quality) was found for Cognitive Behavior Therapy (CBT), multimodal treatment, and Acceptance and Commitment Therapy (ACT). Effects were generally of medium strength and had questionable clinical significance."

Table 3. Continued

Authors, primary	*Country of origin*	*Review primary objectives*	*Search date*	*Index trials included*	*Comparisons affected*	*meta-analysis Y/N*	*narrative synthesis Y/N*	*GRADE Y/N*	*Overall review conclusions*
Shearer et al 2016[41]	Canada	"to update findings of the NPTF and evaluate the effectiveness of psychological interventions for the management of neck pain and associated disorders (NAD) or whiplash associated disorders (WAD)."	To January 2015	6	Physical therapist-provided CBT for Persistent grades I–II NAD, 12-month follow-up	N	Y	N	"We found no clear evidence to support the use of relaxation training, biofeedback, or CBT for reducing pain and improving disability scores in patients with persistent NAD or WAD."
Szeverenyi et al 2018[42]	Hungary, Sweden, USA	"to assess the effectiveness of psychosocial techniques to decrease postoperative pain and improve perioperative clinical care in orthopedic surgery. A systematic review and meta-analysis was performed to evaluate the effects of psychosocial methods among adults undergoing orthopedic surgeries."	To September 2016	8	psychosocial interventions for postoperative pain	Y	N	Y	"The results indicate that psychosocial interventions, especially patient education and relaxation training, may reduce perioperative side effects and improve recovery in patients undergoing orthopedic procedures, but the quality of evidence is generally low."
Vergeld et al 2021[43]	Germany	"to systematically locate and synthesize the current evidence regarding the effectiveness of psychological interventions on fear avoidance beliefs and fear avoidance behavior in patients with CBP."	To October 2019	7	CBT interventions versus active control group or other control groups	N	Y	N	"At this point, the evidence is inconclusive as to which psychological interventions are most effective to treat FAB among people with CBP. Although there is some promising evidence to support CBT, additional research is needed to determine which components of CBT are most effective."

Table 3. Continued

Authors, primary	Country of origin	Review primary objectives	Search date	Index trials included	Comparisons affected	meta-analysis Y/N	narrative synthesis Y/N	GRADE Y/N	Overall review conclusions
Vitoula et al 2018[44]	Greece	"to overview the behavioral approaches that can be used in the management of patients with low back pain. Approaches such as electromyography (EMG) biofeedback, cognitive behavioral therapy, and mindfulness-based stress reduction are discussed as non-pharmacological options in the management of low back pain."	To February 2018	27,9	CBT versus?	N	Y	N	"This systematic review indicates the following key points: 1. Behavioral therapy approaches are effective in patients with LBP particularly in altering pain perception and helping patients to regain their functionality. 2. Treatment outcomes can be improved if the treatments are personalized to individual patients' needs [73,74]. 3. A multidisciplinary approach is the future. Multidisciplinary rehabilitation includes more than just physical treatment. A team approach accounting for several aspects within the bio-psychosocial model is more likely to help individuals with chronic LBP compared to standard care alone. 4. CBT is the type of psychotherapy that has been most studied in patients with LBP. Although most of the other behavioral therapy interventions have been tried in randomized trials in other conditions, more trials of such approaches are needed in patients with LBP. 5. Future research, however, must focus on the improvement of specific outcomes, using not only measures of pain intensity but also using measures of pain acceptance, reduction of medication used, disability, and quality of life to assess efficacy."

Table 3. Continued

Authors, primary	*Country of origin*	*Review primary objectives*	*Search date*	*Index trials included*	*Comparisons affected*	*meta-analysis Y/N*	*narrative synthesis Y/N*	*GRADE Y/N*	*Overall review conclusions*
Williams et al 2020[5]	UK	"To determine the clinical efficacy and safety of psychological interventions for chronic pain in adults (age > 18 years) compared with active controls, or waiting list/ treatment as usual (TAU)."	To April 16, 2020	7,9,11	CBT versus active care	Y	N	Y	"We found sufficient evidence across a large evidence base (59 studies, over 5000 participants) that CBT has small or very small beneficial effects for reducing pain, disability, and distress in chronic pain, but we found insufficient evidence to assess AEs"
Wilson et al 2018[45]	UK	"to determine whether the addition of psychological approaches to physiotherapy is more effective in improving physical functioning and quality of life than physiotherapy alone."	To September 2018	6,7,11	Physiotherapy + psychological versus physiotherapy alone	Y	Y	N	"There is evidence that combining physiotherapy and psychological approaches improves physical function in chronic pain in comparison with physiotherapy alone. Further examination of this field is required to inform changes in practice and to develop treatment methods."
Yang et al 2022[46]	China	"investigated the effectiveness of CBT on pain, disability, fear avoidance, and self-efficacy in patients with CLBP."	To November 20, 2021	7	"CBT versus other therapies"	Y	N	N	"CBT is beneficial in patients with CLBP for improving pain, disability, fear avoidance, and self-efficacy in CLBP patients. Further study is recommended to investigate the long-term benefits of CBT."
Yanyun et al 2021[47]	China	"to systematically review the published evidence to determine whether Pilates exercise training is an efficacious therapy for scoliosis."	To December 2020	10	Pilates versus other	Y	N	N	"Pilates exercise training may reduce the Cobb angle and trunk rotation, relieve pain, increase trunk ROM, and improve QOL for patients with scoliosis. Due to the poor quality of the evidence, however, these results should be interpreted with caution."

Table 4. Summary of Results of Sensitivity Analyses Exploring the Effect of Removing Studies of Interest From Published Meta-Analyses

Group	Outcome (s)	Effect size: number of comparisons	Effect size (SMD) (median IQR) with index trials included	Effect size (SMD) median (IQR) with index trials excluded	Absolute change in effect size (SMD), median (IQR)	% reduction in effect size, median (IQR)	% change in width of 95% confidence intervals, median (IQR)	I2 number of comparisons	Absolute reduction in I2 statistic, median (IQR)
All studies	Pain and Disability	55	-0.6 (–85 to –0.42)	-0.19 (–0.33 to –0.15)	-0.35 (–0.51 to –0.21)	58.3 (39.8–73.7)	-47.6 (–64.6 to –26.5)	44	27.3 (10.3–66.3)
	Pain only	29	-0.52 (–0.73 to –0.4)	-0.19 (–0.29 to –0.15)	-0.35 (–0.49 to –0.19)	62.3 (41.3–73.8)	-47.8 (–65.2 to –25)	24	28.5 (11–74.1)
	Disability only	26	-0.73 (–0.87 to –0.45)	-0.22 (–0.43 to –0.18)	-0.35 (–0.51 to –0.27)	58.2 (40.3–72.8)	-43.5 (–63.4 to –28.5)	20	26.3 (8–43)
Interventions for chronic pain.	Pain and Disability	38	-0.52 (–0.91 to 0.37)	-0.19 (–0.31 to –0.15)	-0.35 (–0.51 to –0.19)	56.9 (41.9–72.8)	-44.5 (–60 to –20)	36	24 (10.3–46.3)
	Pain only	19	-0.48 (–0.83 to –0.34)	-0.19 (–0.30 to –0.15)	-0.34 (–0.44 to –0.17)	50.7 (39–69.9)	-48.9 (–59 to –15)	18	26.5 (11.5–61)
	Disability only	19	-0.63 (–0.95 to –0.4)	-0.19 (–0.33 to –0.16)	-0.35 (–0.55 to –0.21)	58 (46.5–74)	-40 (–60.7 to –23.1)	18	22.5 (8–35.5)
Interventions for postsurgical pain	Pain and disability	17	-0.71 (–0.84 to –0.52)	-0.23 (–0.34 to –0.18)	-0.36 (–0.49 to –0.29)	60.3 (38.3–73.8)	-47.8 (–73.4 to –40.2)	8	88.5 (20.8–95.3)
	Pain only	10	-0.64 (–0.73 to –0.45)	-0.20 (–0.28 to –0.11)	-0.41 (–0.49 to –0.25)	72.9 (47.9–84.1)	-47.7 (–74.9 to –41.5)	6	61 (8.3–95.8)
	Disability only	7	-0.84 (–0.85 to –0.69)	-0.31 (–0.57 to –0.22)	-0.35 (–0.42 to –0.3)	58.3 (36.2–61.8)	-53 (–72.1 to –37.1)	2	88.5 (87.3–89.8)

Abbreviations: IQR, Inter Quartlie Range; SMD, Standardized Mean Difference.

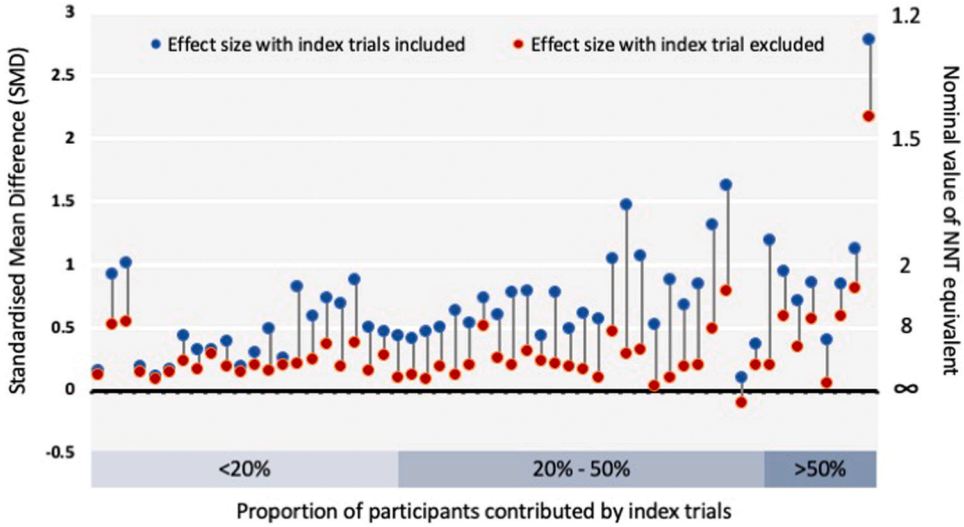

Figure 3. The effect of excluding index trials from the included meta-analyses. Footnotes: NNT estimated from the SMD using the approach described by Faraone.[66] Abbreviation: SMD, Standardized Mean Difference.

Impact on Overall Review Conclusions

Twenty (63%) of the included reviews reported broadly positive conclusions regarding the effectiveness of the intervention of interest, 6 (19%) reported cautiously positive conclusions which were qualified by issues of the quality of the evidence, or the size of treatments, and 6 (19%) came to equivocal or negative conclusions. While we could not accurately predict review authors' interpretations to analyses after the removal of the index trials, we might expect these findings to impact either their overall conclusions and/or the certainty around those conclusions for many of those reviews.

Impact on Clinical Practice Guidelines

We included CPGs from Belgium,[57] Canada,[49,51] Finland,[50] Japan,[53] the Netherlands,[48] Russia,[56] the UK,[55] and the USA.[52,54] Seven CPGs[48,52–57] were focused on the management of low back pain, one on the management of chronic nonmalignant pain,[51] one on the management of whiplash-associated disorder and neck pain-associated disorders,[49] and 1 on neck pain in adults.[50] Table 5 presents the characteristics of the included CPGs. 4 CPGs[48,49,55,57] used GRADE to evaluate the certainty of evidence.

Nine CPGs presented narrative syntheses for comparisons that included index trials and one[53] conducted a de novo meta-analysis. The interventions of interest were described as multimodal, multidisciplinary, or biopsychosocial,[49,51,55–57] CBT combined with exercise,[54] behavioral treatment,[48] cognitive therapy,[50] or general exercise.[52] Table 6 summarizes the CPG analyses that included the index trials. All but one made positive recommendations for interventions for which index trials had informed the synthesis. No guideline raised any concerns regarding the veracity of the index trials.

In the Academy of Orthopedic Physical Therapy of the American Physical Therapy Association guideline,[52] Monticone 2016a was 1 of 4 RCTs included in the evaluation of "general exercise" for low back pain. Of the 4 RCTs in this category, Monticone 2016a was the only study reported to show benefit (of CBT added to general group-based exercise vs exercise alone). Recommendations were derived from evidence from a wider range of exercise categories but the guideline did specifically recommend that physical therapists should use "general exercise".

In the Canadian Agency for Drugs and Technologies in Health guideline,[51] 2 systematic reviews and 2 additional RCTs, including Monticone 2017, informed the evaluation of the clinical effectiveness of multidisciplinary treatment programmes for persistent nonmalignant pain. Both systematic reviews were reported to demonstrate benefit, one in the short term but not long term, and the other at long-term follow-up. The number of trials and participants in those reviews was not reported. Of the newly included trials, only Monticone 2017 was reported to show a benefit of a multidisciplinary programme, including at 12-month follow-up. The guideline concluded that multidisciplinary management of chronic nonmalignant pain was associated with significant improvements in pain intensity, and may be associated with significant improvements in quality of life and function.

The Canadian Chiropractic Association and the Canadian Federation of Chiropractic Regulatory and Educational Accrediting Boards guideline[49] included 3 RCTs in its evidence evaluation of multimodal care versus continued practitioner care for neck pain-associated disorder, of which Monticone 2012 was one. The guideline reported that Monticone 2012 did not demonstrate clinically significant effects at 1-year follow-up, but made a positive recommendation for multimodal care on the basis of the other 2 trials, which were reported to show benefit.

The Finnish medical association[50] guideline included the 3 RCTS and one nonrandomized controlled trial in

Table 5. Characteristics of Clinical Practice Guidelines That Included Index Trials

Organization	*Title primary*	*Year published*	*Country of origin*	*Review Objectives*	*Search dates*	*Index trials included*	*Interventions and /or Comparisons affected*	*meta-analysis? Y/N*	*narrative synthesis Y/N*	*GRADE Y/N*
Academy of Orthopedic Physical Therapy of the American Physical Therapy Association (AAOPT)[52]	Interventions for the Management of Acute and Chronic Low Back Pain: Revision 2021	2021	USA	CPG for acute and chronic LBP	To June 25, 2020	Monticone 2016a	General exercise to other exercise training interventions	N	Y	N
Belgian Health Care Knowledge Centre (KCE)[57]	Low back pain and radicular pain: assessment and management	2017	Belgium	The guidelines provide recommendations based on current scientific evidence for the evaluation and management of low back pain and radicular Pain in the adult population. Clinicians.	to Dec 2015	Monticone 2013, 2016a	Multidisciplinary biopsychosocial treatment programs	N	Y	Y
Canadian Agency for Drugs and Technologies in Health (CADTH)[51]	Multidisciplinary treatment programs for patients with chronic non-malignant Pain: a review of clinical effectiveness, cost-effectiveness, and guidelines: an update	2019	Canada	to review the comparative clinical effectiveness, cost-effectiveness, and evidence-based guidelines regarding the use of multidisciplinary treatment programs for patients with chronic, non-malignant Pain in outpatient settings	To April 2019	Monticone 2017	Multidisciplinary treatment programs	N	Y	N
Canadian Chiropractic Association and the Canadian Federation of Chiropractic Regulatory and Educational Accrediting Boards (CCA/ CFC)[49]	The Treatment of Neck Pain-Associated Disorders and Whiplash-Associated Disorders: A Clinical Practice Guideline	2016	Canada	to determine which treatments or combinations of treatments are more effective for managing NAD and WAD.	To December 24, 2015	Monticone 2012	Multimodal care vs continued practitioner care for persistent grades I to III NAD?	N	Y	Y
Dutch General Practitioners Association. Nederlands Huisarten Genootschap (NHG)[48]	Non-specific low back pain (M54)	2017	Netherlands	CPG for the diagnosis and management of non-specific low back Pain.	To June 2014	Monticone 2013	Behavioral treatment and exercise therapy vs. exercise therapy	N	Y	Y

Table 5. Continued

Organization	Title primary	Year published	Country of origin	Review Objectives	Search dates	Index trials included	Interventions and /or Comparisons affected	meta-analysis? Y/N	narrative synthesis Y/N	GRADE Y/N
Finnish Medical Association, Association of Physical Medicine and Rehabilitation Fennia (FMA)[50]	Neck Pain	2017	Finland	CPG for the management of neck Pain	NR	Monticone 2012	Cognitive therapy	N	Y	N
Japanese Orthopedic Association (JOA)[53]	Japanese Orthopedic Association (JOA) clinical practice guidelines on the management of lumbar spinal stenosis, 2021 - Secondary publication	2022	Japan	biopsychosocial rehabilitation for patients with chronic low back Pain.	2008–2019	Monticone 2014B	Effects of postoperative physiotherapy 1-year after surgery	Y	Y	N
National Institute of Health and Care Excellence (NICE)[55]	NG59 low back pain and sciatica in over 16 s: assessment and management	2016	UK	Clinical guidelines for the assessment and management of low back pain, with or without sciatica	to Dec 2015	Monticone 2013, 2016a	Multidisciplinary biopsychosocial treatment programs	N	Y	Y
North American Spine Society (NASS)[54]	Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of low back Pain	2020	USA	To provide an evidence-based educational tool to assist spine specialists when making clinical decisions for adult patients with nonspecific low back Pain.	NR	Monticone 2013	Exercise therapy alone versus exercise with cognitive behavioral therapy (CBT) Interventions that address fear-avoidance behaviors	N	Y	N
Russian Society for the Study of Pain (RSSP)[56]	Chronic nonspecific (musculoskeletal) low back Pain. Guidelines of the Russian Society for the Study of Pain (RSSP)	2019	Russia	Chronic nonspecific (musculoskeletal) low back Pain. Guidelines of the Russian Society for the Study of Pain	NR	Monticone 2013. 2016A	Multidisciplinary biopsychosocial treatment programs	N	Y	N

Abbreviation: NR, not reported.
Recommendations based on evidence that included a study of interest.

Table 6. Summary of Guideline Analyses That Included Index Trials

CPG	Affected research question or comparison	Index trials included	Synthesis summary	Recommendation or conclusions
Academy of Orthopedic Physical Therapy of the American Physical Therapy Association (AAOPT)[52]	How effective is general exercise compared to other exercise training interventions?	9	4 trials of exercise training included. Combined N not reported. N contributed by the trial of interest= 150	"Physical therapists should use exercise training interventions, including trunk muscle strengthening and endurance, multimodal exercise interventions, specific trunk muscle activation exercise, aerobic exercise, aquatic exercise, and general exercise, for patients with chronic LBP."
Belgian Health Care Knowledge Centre (KCE)[57]	What is the effectiveness biopsychosocial treatment programs for low back Pain with or without sciatica?	7,9	3 trials of multidisciplinary biopsychosocial treatment programs included. Combined N = 361 N contributed by index trials = 240	"Consider a multidisciplinary rehabilitation programme, which combines a physical and a psychological component, incorporating a cognitive behavioral approach, and which takes into account a person's specific needs and capabilities, for people with persistent low back pain or radicular Pain: or when they have psychosocial obstacles to recovery or when previous evidence-based management has not been effective."
Canadian Agency for Drugs and Technologies in Health (CADTH)[51]	To review the comparative clinical effectiveness, cost-effectiveness, and evidence-based guidelines regarding the use of multidisciplinary treatment programs for patients with chronic, nonmalignant Pain in outpatient settings	11	2 Systematic reviews and 2 RCTS included. RCTS combined N = 329. N contributed by the trial of interest= 170	"Overall, findings from the included studies suggested that the multidisciplinary management of chronic non-malignant Pain was associated with significant improvements in Pain intensity, and may be associated with significant improvements in quality of life and function."
Canadian Chiropractic Association and the Canadian Federation of Chiropractic Regulatory and Educational Accrediting Boards (CCA/ CFC)[49]	Should multimodal care versus continued practitioner care be used for persistent grades I to III NAD?	6	2 RCTs, combined N = 357. N contributed by trial of interest = 90. Trial of interest not considered to show evidence of clinically important effects. "The addition of a cognitive behavioral treatment did not provide greater outcomes than multimodal care alone."	"For patients presenting with persistent neck Pain grades I to III, we suggest clinicians offer multimodal care* and/ or practitioner advice† based on patient preference. (Weak recommendation, low-quality evidence)"
Dutch General Practitioners Association[48]	Behavioral treatment and exercise therapy versus exercise therapy	7	1 systematic review with 30 RCTs (combined N = 3438) + 5 additional RCTs (combined N = 889). N contributed by index trials = 90	"It is unclear whether there are clinically relevant benefits of cognitive behavioral treatment over waiting list or standard treatment in patients with chronic non-specific low back Pain."
Finnish Medical Association, Association of Physical Medicine and Rehabilitation Fennia[50]	The effectiveness of cognitive therapies for neck Pain	6	3 RCTS AND 1 controlled clinical trial (combined N = 813) N contributed by index trials = 80	"Cognitive therapy may be effective in the treatment of neck pain, but there is no convincing research evidence of the effectiveness of such therapy".
Japanese Orthopedic Association (JOA)[53]		8	Meta-analyses: Physiotherapy for leg Pain at one year. 5 RCTs, combined N = 595	"Physiotherapy for patients undergoing surgical treatment for LSS is effective for alleviating pain

Table 6. Continued

CPG	Affected research question or comparison	Index trials included	Synthesis summary	Recommendation or conclusions
	What are the effects of postoperative physiotherapy 1-year after surgery for spinal stenosis?		N contributed by index trials = 117 Physiotherapy for back pain at one year. 4 RCTs, combined N = 473 N contributed by index trials = 117	and improving ADL and QOL 3 months after surgery and adverse events are rare. Thus, postoperative physiotherapy can be considered useful."
National Institute of Health and Care Excellence (NICE)[55]	What is the effectiveness biopsychosocial treatment programs for low back Pain with or without sciatica?	7,9	3 RCTs of multidisciplinary biopsychosocial treatment programs included. Combined N = 361 N contributed by index trials = 240	"Consider a multidisciplinary rehabilitation programme, which combines a physical and a psychological component, incorporating a cognitive behavioral approach, and which takes into account a person's specific needs and capabilities, for people with persistent low back pain or radicular Pain: or when they have psychosocial obstacles to recovery or when previous evidence-based management has not been effective."
North American Spine Society (NASS)[54]	What are the outcomes, including duration of Pain, intensity of Pain, functional outcomes, and return-to-work status, for exercise therapy alone versus exercise with cognitive behavioral therapy (CBT)?	7	8 RCTs of the addition of CBT to exercise. Combined N = 913 N contributed by index trials = 90 4 RCTs of interventions to reduce fear avoidance. Combined N = 287 N contributed by index trials = 90	"There is conflicting evidence that addition of CBT to an exercise program results in significant improvement in Pain and function compared with exercise alone in patients with chronic low back Pain." "Treatments targeting fear avoidance combined with physical therapy are recommended compared to physical therapy alone to improve low back Pain in the first 6 months."
Russian Society for the Study of Pain (RSSP)[56]	What is the effectiveness of Multidisciplinary biopsychosocial treatment programs?	7,9	23 studies referenced in the evidence summary for multidisciplinary biopsychosocial treatment programs. Combined N or further details not reported. N contributed by index trials = 240	"Multidisciplinary programs, which include physical therapy (therapeutic exercises), psychological methods (primarily cognitive-behavioral therapy), and educational conversations/lectures (schools) for patients are recommended for the treatment of chronic LBP. The use of multidisciplinary programs allows for improving the main indicators of the patient's condition: intensity of Pain, functional activity, professional activity, psychological and physical quality of life."

its evaluation of the effectiveness of cognitive therapies for neck pain (combined n = 813). These included the index trial Monticone 2012 (n = 80). The guideline concluded that while cognitive therapy may be effective, there was no convincing research evidence to that effect, but reported no further detail.

The Japanese Orthopedic association[53] guideline included Monticone 2014 in its synthesis of the effectiveness of postoperative physiotherapy after surgery for spinal stenosis. It conducted meta-analyses for back pain intensity and leg pain intensity, activities of daily living (ADL), health-related quality of life, and general health at 1-year postsurgery that all included Monticone 2014. We conducted sensitivity analyses where Monticone 2014 was removed from these meta-analyses. The results are presented in Supplementary Table 5. Our sensitivity analyses resulted in smaller point estimates of effect for all analyses, with greater precision for most, and reduced heterogeneity in 3 out of 4 analyses. The guideline concluded that physiotherapy was effective at alleviating pain and improving activities of daily living and QoL, and could therefore be considered useful.

The Nederlands Huisarten Genootschap NHG[48] guideline included a systematic review of 30 RCTS (combined n = 3438) and a further 5 RCTs (combined n = 889) in its evaluation of behavioral therapies for low back pain which included the index trial Monticone 2013 (n = 90). When reporting on the index trial, the guideline stated that "*the presentation of the results made it impossible to assess the clinical relevance*" of the treatment effect but did not elaborate further. Overall, it concluded that it was unclear whether there were clinically relevant benefits of cognitive behavioral treatment over the waiting list or standard treatment in patients with chronic nonspecific low back pain (low to very low overall quality of evidence).

The UK National Institute of Health and Care Excellence (NICE)[55] guidance in pain and the Belgian Health Care Knowledge Centre (KCE) guidance will be considered together as the KCE guidance was based on the evidence synthesis conducted by NICE. NICE included 2 studies of interest (Monticone 2013, 2016a) in the evaluation of "*Multidisciplinary biopsychosocial rehabilitation (MBR) programmes for back pain*". The guideline found the evidence for MBR programmes to be mixed, with clinical benefits seen for some comparisons, but also many instances where no benefit was observed and a few where the comparator was favored over MBR. The guideline reports that the evidence that informed the guidelines for this intervention came "primarily" from the RCTs (combined n = 361), including Monticone 2013 and 2016a (combined n = 240). These trials all reported benefits for MBR programmes and so NICE undertook de novo threshold analyses for the cost-effectiveness of MBR programmes based on the results of these trials and concluded that the interventions in both the Monticone 2013 and 2016a studies were likely to be cost-effective. NICE recommended that clinicians should "*Consider a multidisciplinary rehabilitation programme … for people with persistent low back pain or radicular pain: or when they have psychosocial obstacles to recovery or when previous evidence-based management has not been effective*."

The North American Spine Society NASS[54] guideline included Monticone 2013 in its evaluation of the effectiveness of interventions that address fear and avoidance. Overall, they included 4 RCTs for this comparison (combined n = 287), of which Monticone 2013 contributed 90 participants. Three of 4 studies, including Monticone 2013, reported benefits on fear and avoidance outcomes, and the guideline recommended such interventions. A separate comparison of the effectiveness of adding CBT to an exercise programme included 8 RCTs (combined n = 913) of which Monticone 2013 was 1 (90 participants). The guideline reported conflicting evidence for the addition of CBT and did not make a specific recommendation for its use.

In its evaluation of the effectiveness of multidisciplinary biopsychosocial treatment programmes, the Russian Society for the Study of Pain guideline[56] included 23 studies in their evidence summary. It was not clear whether these were all RCTs, and the sample size of these studies was not provided. The guideline included Monticone 2013 and 2016a, which together randomized 240 participants. Little detail was provided on how each study informed the guideline recommendation but the guideline recommended multidisciplinary interventions for chronic low back pain.

Discussion

We previously identified concerns regarding the trustworthiness of this cohort of trials[1] relating to aspects of research governance, data anomalies, and implausible results. In our current study, we found that these index trials have been included in 32 systematic reviews and have had important impacts on the results and conclusions of the majority of those. The inclusion of the index trials has exaggerated the size of estimated treatment effects, increased inconsistency in meta-analyses, and altered the precision of meta-analyses. In many cases, the exclusion of index trials changed the pooled effects of meta-analyses from moderate-to-large to small or very small effect sizes. These new effect estimates are of questionable clinical significance and, in some cases, excluding index trials shifted effects from statistically significant to nonsignificant. Index trials have also influenced reviews undertaking narrative syntheses. While the impact on narrative syntheses is harder to quantify, it is reasonable to conclude that index trials weighted the conclusions of many of those reviews in an unduly positive direction.

We identified a number of CPGs from a range of countries and organizations that included at least one of the index trials and used them to formulate their recommendations. All CPGs made positive recommendations for the interventions for which index trials informed the syntheses. Due to the varied approaches to reporting in CPGs and the dominance of

narrative approaches to syntheses, it is often not possible to ascertain the specific contribution of index trials to their conclusions and recommendations. In most included CPGs, it is reasonable to infer that the positive reported findings of the index trials contributed to recommendations that favored psychological or multimodal therapies. In specific examples, it is clear that the index trials were crucial to such clinical recommendations. The NICE 2016 guideline[55] clearly shows that 2 of the index trials[7,9] were included in the 3 trials whose evidence was used for a de novo economic analysis that drove a recommendation for multidisciplinary biopsychosocial rehabilitation for low back pain. It is not unreasonable to speculate that without the index trials, such a recommendation would not have been considered appropriate. That the evidence in the NICE guideline[55] was directly used in the formulation of the Belgian (KCE)[57] guideline further extends that impact.

There is a parallel in the field of hip fracture, an exploration of the impact of a cohort of trials from a different single lead author, that were affected by research misconduct, similarly found that those trials significantly distorted the findings of reviews and clinical practice guidelines[58].

As we have recently demonstrated,[1] the index trials are extreme outliers in terms of the size of reported effects of psychological therapies, specifically CBT. Just under half of the included reviews specifically referred to the outlier status of index trials and, of those, 4 conducted and reported sensitivity analyses that excluded the index trials. Only 1 (Cochrane) review excluded the index trials from the primary analyses[5] on the basis that the observed heterogeneity was not satisfactorily explained. Other reviews did not comment further on the observed heterogeneity and a number speculated that it might be the result of specific intervention characteristics. Importantly, no review or CPG specifically raised concerns regarding the trustworthiness of the index trials.

This last observation is unsurprising. Screening for aspects of trustworthiness of studies has not yet become routine practice in evidence synthesis and indeed the development and validation of tools for this purpose is in its infancy. While systematic reviews routinely assess the risk of bias, trustworthiness screening has a distinct and broader scope. It is noteworthy that the risk of bias assessments for the index trials was unremarkable[1] for this field of study. As we have previously argued,[1] systematic reviewers and guideline developers need to attend to the possibility that a range of other factors, including error, poor research governance, and/or misconduct, may affect identified studies and to develop and adopt approaches to this. Tools are beginning to emerge[59–62] that offer some structure to this task for prospective systematic reviewers, though further evaluation of their validity and performance is needed. Our results provide strong support for the argument for such screening. Using these tools, systematic reviewers might move to a process in which trustworthiness is not assumed, and where studies that do not clearly meet a threshold of trustworthiness, through pre-registration, evidence of good research governance, and methodological and data transparency, are not included in the synthesis of evidence and cannot influence review conclusions.

Regardless of formal screening for trustworthiness, we propose that reviewers routinely identify and carefully scrutinize studies with divergent results in their evidence syntheses and, where possible, seek explanations from the authors of those studies. Reluctance to do this can be driven by a commitment to follow a protocol and the additional resource burden to reviewers, editors, and guideline developers. However, to fail to implement any approach presents a risk of the uncritical inclusion of misleading data.

Where an evidence base is dominated by small trials with generally modest effects, as is the case with psychological therapies for chronic pain,[5] the inclusion of untrustworthy data can seriously impact results.[63] This places conclusions regarding the effectiveness and the decisions of guideline developers in a marginal space where subtle differences in interpretation can lead to meaningful differences in recommendations.[64] The introduction of trials with unremarkable risk of bias profiles[1] but very large treatment effects can have a particularly large impact, both on pooled effect sizes and on how that body of evidence is interpreted in research and practice. Our study provides clear evidence of this.

We have used a language of trustworthiness as it accurately reflects the process of assigning a judgment of whether one has trust in the veracity of the findings. The benefits of this approach are that we are clear about where that judgment lies and the basis of that judgment. This is in line with a similar approach in GRADE in moving away from attempting an objective rating of quality to a subjective judgment of certainty of evidence.[65] The risk of this approach is that it introduces terminology that could be misconstrued as a judgment of researcher behavior or intent. As this field develops, and new tools and methods are introduced, we would also expect the terminology to be also further developed.

Our study has some specific strengths. We followed a publicly available protocol, used tools to identify citations for the index trials, and screened the results of those searches with independent reviewers. We have developed a novel multidimensional index to classify the impact that index trials have had on meta-analyses and look forward to other researchers scrutinizing and refining it. There are also some limitations. Data extraction was conducted by a single reviewer, though all analyses were checked by a second reviewer. The undertaking of sensitivity analyses based on the data published in the identified reviews, rather than in the original trials, risks replicating errors contained in those reviews, but the principle aim of our reanalyses was to ascertain the impact that the index trials had on the reported analyses in those reviews, rather than to estimate the effects of interventions. The use of the I^2 statistic when considering consistency and heterogeneity in meta-analyses has been criticized, as it is not

a direct measure of heterogeneity but rather of the percentage of variability in effect estimates in a meta-analysis due to between-trial heterogeneity rather than chance. However, for this study, between-trial heterogeneity driven by the inclusion of the index trials was most relevant to our purposes. As several included reviews and CPGs did not implement the GRADE approach to evaluating certainty and performing de novo GRADE judgments from the available information reported in reviews and CPGs would have not been possible in many instances we did not evaluate the impact on GRADE judgments.

The development and application of the Impact Index was not planned in our original protocol but developed through internal consensus within the team; it has not been formally validated. We propose that it has reasonable face validity. We are aware that including aspects of statistical significance in the application of the index will be controversial. However, in practice, changes in the statistical significance of meta-analyses frequently influence the conclusions of reviews and, indeed, the interpretation of those reviews by their readers. There remains a need to formally evaluate the impact index we developed here and to more rigorously test its assumptions.

Our study represents the impact of the index studies at the point at which the searches were conducted. At the time of writing, we have raised concerns regarding these trials with the editors of their host journals: 3 of the trials have been formally retracted and some investigations are proceeding. Nevertheless, it is likely that some of these trials, including potentially those that have been retracted, will be included in future systematic reviews and CPGs.

Our findings have important implications for the application of evidence-based healthcare. RCTs and systematic reviews of RCTs are routinely held up as the reference standard of evidence for ascertaining the effectiveness of interventions and for underpinning clinical recommendations. We have shown here how this cohort of studies has led to substantial impacts on both the results of systematic reviews and the recommendations of CPGs, contributing to overly positive conclusions regarding the effectiveness of adding CBT to physical rehabilitation for spinal pain, with subsequent impacts on clinical decisions.

Transparency Declaration

The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Data Sharing Statement

The full data underpinning the analysis of the impact of studies on published meta-analyses are available via Figshare: https://doi.org/10.17633/rd.brunel.21427995.v1

Disclosures

This study was not supported by any external funding. Michael Wewege was supported by a Postgraduate Scholarship from the National Health and Medical Research Council of Australia and a PhD Supplementary Scholarship from Neuroscience Research Australia. Christopher Eccleston undertook consultancy for Orion Pharma and Reckitt on the psychology of pain. He is a member of the UK Pain Parliamentary Campaign Group with a focus on raising awareness in government on the needs of pain patients. Amanda Williams undertook consultancy for Reckitt on the psychology of pain. NOC was a member of the Guideline Development Group for the 2016 NICE guidance for low back pain and Sciatica (NG59) which is included in this study. AW, CE, EF, and LH were authors on one included review.[5] They were not involved in any data extraction or analysis related to that review for this project.

Contributor and guarantor information

NO'C is the guarantor and accepts overall responsibility for the work.

NO'C, CE, GS, EF, AW and LH developed the original concept for the study and developed the protocol.

NO'C conducted the searches; NO'C and LH screened the searches.

NO'C and MAW extracted and verified the extracted data.

NO'C and MAW conducted the data analyses.

All authors contributed to drafting of the manuscript.

The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpain.2023.07.003.

References

1. O'Connell NE, Moore RA, Stewart G, *et al.* Investigating the veracity of a sample of divergent published trial data in spinal pain. Pain 164(1):72-83, 2022. https://doi.org/10.1097/j.pain.0000000000002659

2. Bordewijk EM, Li W, van Eekelen R, *et al.* Methods to assess research misconduct in health-related research: A scoping review. J Clin Epidemiol 136:189-202, 2021. https://doi.org/10.1016/j.jclinepi.2021.05.012

3. Bolland MJ, Grey A, Avenell A, Klein AA: Correcting the scientific record—A broken system? Account Res 28(5):265-279, 2021. https://doi.org/10.1080/08989621.2020.1852938

4. Kataoka Y, Banno M, Tsujimoto Y, *et al.* Retracted randomized controlled trials were cited and not corrected in

systematic reviews and clinical practice guidelines. J Clin Epidemiol 150:90-97, 2022. https://doi.org/10.1016/j.jclinepi.2022.06.015

5. Williams AC de C, Fisher E, Hearn L, Eccleston C: Psychological therapies for the management of chronic pain (excluding headache) in adults. Cochrane Database Syst Rev 8(8):CD007407. https://doi.org/10.1002/14651858.CD007407.pub4, 2020.

6. Monticone M, Baiardi P, Vanti C, *et al.* Chronic neck pain and treatment of cognitive and behavioural factors: Results of a randomised controlled clinical trial. Eur Spine J 21(8):1558-1566, 2012. https://doi.org/10.1007/s00586-012-2287-y

7. Monticone M, Ferrante S, Rocca B, Baiardi P, Dal Farra F, Foti C: Effect of a long-lasting multidisciplinary program on disability and fear-avoidance behaviors in patients with chronic low back pain: Results of a randomized controlled trial. Clin J Pain 29(11):929-938, 2013. https://doi.org/10.1097/AJP.0b013e31827fef7e

8. Monticone M, Ferrante S, Teli M, *et al.* Management of catastrophising and kinesiophobia improves rehabilitation after fusion for lumbar spondylolisthesis and stenosis. A randomised controlled trial. Eur Spine J 23(1):87-95, 2014. https://doi.org/10.1007/s00586-013-2889-z

9. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Foti C: Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain. Eur J Pain 20:541-551, 2016. https://doi.org/10.1002/ejp.756

10. Monticone M, Ambrosini E, Cazzaniga D, *et al.* Adults with idiopathic scoliosis improve disability after motor and cognitive rehabilitation: Results of a randomised controlled trial. Eur Spine J 25(10):3120-3129, 2016. https://doi.org/10.1007/s00586-016-4528-y

11. Monticone M, Ambrosini E, Rocca B, *et al.* Group-based multimodal exercises integrated with cognitive-behavioural therapy improve disability, pain and quality of life of subjects with chronic neck pain: A randomized controlled trial with one-year follow-up. Clin Rehabil 31(6):742-752, 2017. https://doi.org/10.1177/0269215516651979

12. Monticone M, Ambrosini E, Rocca B, *et al.* Multimodal exercises integrated with cognitive-behavioural therapy improve disability of patients with failed back surgery syndrome: A randomized controlled trial with one-year follow-up. Disabil Rehabil 27:1-8, 2020. https://doi.org/10.1080/09638288.2020.1863480

13. Monticone M, Ambrosini E, Portoghese I, Rocca B: Multidisciplinary program based on early management of psychological factors reduces disability of patients with subacute low back pain. Results of a randomised controlled study with one year follow-up. Eur J Phys Rehabil Med 57(6):959-967, 2021. https://doi.org/10.23736/S1973-9087.21.06696-X

14. Haddaway N.R., Grainger M.J., Gray C.T. citationchaser: An R package and Shiny app for forward and backward citations chasing in academic searching; 2021. doi:⟨10.5281/zenodo.4543513⟩.

15. Review Manager (RevMan). Version 5.4. The Cochrane Collaboration; 2020.

16. Rücker G, Schwarzer G, Carpenter JR: Undue reliance on I^2 in assessing heterogeneity may mislead. BMC Med Res Methodol 8:79, 2008. https://doi.org/10.1186/1471-2288-8-79

17. Alanazi MH, Parent EC, Dennett E: Effect of stabilization exercise on back pain, disability and quality of life in adults with scoliosis: A systematic review. Eur J Phys Rehab Med 54(5):647-653, 2017. https://doi.org/10.23736/S1973-9087.17.05062-6

18. Barbari V, Storari L, Ciuro A, Testa M: Effectiveness of communicative and educative strategies in chronic low back pain patients: A systematic review. Patient Educ Couns 103(5):908-929, 2019. https://doi.org/10.1016/j.pec.2019.11.031

19. Bogaert L, Thys T, Depreitere B, *et al.* Rehabilitation to improve outcomes of lumbar fusion surgery: A systematic review with meta-analysis. Eur Spine J 31(6):1525-1545, 2022. https://doi.org/10.1007/s00586-022-07158-2

20. Casey M, Smart KM, Segurado R, Doody C: Multidisciplinary-based Rehabilitation (MBR) compared with active physical interventions for pain and disability in adults with chronic pain: A systematic review and meta-analysis. Clin J Pain 36(11):874-886, 2020. https://doi.org/10.1097/AJP.0000000000000871

21. Cheng JOS, Cheng S: Effectiveness of physical and cognitive-behavioural intervention programmes for chronic musculoskeletal pain in adults: A systematic review and meta-analysis of randomised controlled trials. PLoS One 14(10):e0223367. https://doi.org/10.1371/journal.pone.0223367, 2019.

22. Fadli A, Prasetya H, Kristyanto A: Effectivity biopsychosocial intervention with cognitive behavioral therapy and exercise therapy program in chronic low back pain: Meta-analysis. Indonesian J Med 6(2):177-193, 2021. https://doi.org/10.26911/theijmed.2021.06.02.07

23. Greenwood J, McGregor A, Jones F, Hurley M: Rehabilitation following lumbar fusion surgery: A randomised controlled feasibility study. Eur Spine J 28(4):735-744, 2019. https://doi.org/10.1007/s00586-019-05913-6

24. Guerrero AVS, Maujean A, Campbell L, Sterling M: A systematic review and meta-analysis of the effectiveness of psychological interventions delivered by physiotherapists on pain, disability and psychological outcomes in musculoskeletal pain conditions. Clin J Pain 34(9):838-857, 2018. https://doi.org/10.1097/AJP.0000000000000601

25. Hajihasani A, Rouhani M, Salavati M, Hedayati R, Kahlaee AH: The influence of cognitive behavioral therapy on pain, quality of life, and depression in patients receiving physical therapy for chronic low back pain: A systematic review. PM & R 11(2):167-176, 2019. https://doi.org/10.1016/j.pmrj.2018.09.029

26. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder M,W: Exercise therapy for chronic low back pain. Cochrane Database Syst Rev 9(9):CD009790. https://doi.org/10.1002/14651858.CD009790.pub2, 2021.

27. Ho EK, Chen L, Simic M, *et al.* Psychological interventions for chronic, non-specific low back pain: Systematic review with network meta-analysis. BMJ 376:e067718. https://doi.org/10.1136/bmj-2021-067718, 2022.

28. İnce B: Systematic review of the comparative effectiveness of cognitive-behavioural therapies for chronic pain. J Cognit Behav Psychother Res 9(3):248-259, 2020. https://doi.org/10.5455/JCBPR.64324

29. Kamper SJ, Apeldoorn AT, Chiarotto A, *et al.* Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and

meta-analysis. BMJ 350:h444. https://doi.org/10.1136/bmj.h444, 2015.

30. Knoerl R, Lavoie Smith EM, Weisberg J: Chronic pain and cognitive behavioral therapy: An integrative review. West J Nurs Res 38(5):596-628, 2016. https://doi.org/10.1177/0193945915615869

31. Martinez-Calderon J, Flores-Cortes M, Morales-Asencio JM, Luque-Suarez A: Conservative interventions reduce fear in individuals with chronic low back pain: A systematic review. Arch Phys Med Rehabil 101(2):329-358, 2019. https://doi.org/10.1016/j.apmr.2019.08.470

32. Monticone M, Ambrosini E, Cedraschi C, *et al.* Cognitive-behavioral treatment for subacute and chronic neck pain. Spine 40(19):1495-1504, 2015. https://doi.org/10.1097/BRS.0000000000001052

33. Nadinda PG, van Ryckeghem DML, Peters ML: Can perioperative psychological interventions decrease the risk of postsurgical pain and disability? A systematic review and meta-analysis of randomized controlled trials. Pain 163(7):1254-1273, 2022. https://doi.org/10.1097/j.pain.0000000000002521

34. Nicholls JL, Azam MA, Burns LC, *et al.* Psychological treatments for the management of postsurgical pain: A systematic review of randomized controlled trials. Patient Relat Outcome Meas 9:49-64, 2018. https://doi.org/10.2147/PROM.S121251

35. O'Keeffe M, Purtill H, Kennedy N, *et al.* Comparative effectiveness of conservative interventions for nonspecific chronic spinal pain: Physical, behavioral/psychologically informed, or combined? A systematic review and meta-analysis. J Pain 17(7):755-774, 2016. https://doi.org/10.1016/j.jpain.2016.01.473

36. Özden F: The effectiveness of physical exercise after lumbar fusion surgery: A systematic review and meta-analysis. World Neurosurg 163:396-412, 2022. https://doi.org/10.1016/j.wneu.2022.03.143

37. Parrish JM, Jenkins NW, Parrish MS, *et al.* The influence of cognitive behavioral therapy on lumbar spine surgery outcomes: A systematic review and meta-analysis. Eur Spine J 30(5):1365-1379, 2021. https://doi.org/10.1007/s00586-021-06747-x

38. Petrucci G, Papalia GF, Russo F, *et al.* Psychological approaches for the integrative care of chronic low back pain: A systematic review and metanalysis. Int J Environ Res Public Health 19(1):60, 2021. https://doi.org/10.3390/ijerph19010060

39. Richmond H, Hall AM, Copsey B, *et al.* The effectiveness of cognitive behavioural treatment for non-specific low back pain: A systematic review and meta-analysis. PLoS One 10(8):e0134192. https://doi.org/10.1371/journal.pone.0134192, 2015.

40. Schütze R, Rees CS, Smith A, Slater H, Campbell JM, O'Sullivan P: How can we best reduce pain catastrophizing in adults with chronic noncancer pain? A systematic review and meta-analysis. J Pain 19(3):233-256, 2017. https://doi.org/10.1016/j.jpain.2017.09.010

41. Shearer HM, Carroll LJ, Wong JJ, *et al.* Are psychological interventions effective for the management of neck pain and whiplash-associated disorders? A systematic review by the Ontario Protocol for Traffic Injury Management (OPTIMa) collaboration. Spine J 16(12):1566-1581, 2015. https://doi.org/10.1016/j.spinee.2015.08.011

42. Szeverenyi C, Kekecs Z, Johnson A, Elkins G, Csernatony Z, Varga K: The use of adjunct psychosocial interventions can decrease postoperative pain and improve the quality of clinical care in orthopedic surgery: A systematic review and meta-analysis of randomized controlled trials. J Pain 19(11):1231-1252, 2018. https://doi.org/10.1016/j.jpain.2018.05.006

43. Vergeld V, Ginis KAM, Jenks AD: Psychological interventions for reducing fear avoidance beliefs among people with chronic back pain. Rehabil Psychol 66(4):386-403, 2021. https://doi.org/10.1037/rep0000394

44. Vitoula K, Venneri A, Varrassi G, *et al.* Behavioral therapy approaches for the management of low back pain: An up-to-date systematic review. Pain Ther 7(1):1-12, 2018. https://doi.org/10.1007/s40122-018-0099-4

45. Wilson S, Cramp F: Combining a psychological intervention with physiotherapy: A systematic review to determine the effect on physical function and quality of life for adults with chronic pain. Phys Ther Rev 23(3):214-226, 2018. https://doi.org/10.1080/10833196.2018.1483550

46. Yang J, Lo WLA, Zheng F, Cheng X, Yu Q, Wang C: Evaluation of cognitive behavioral therapy on improving pain, fear avoidance, and self-efficacy in patients with chronic low back pain: A systematic review and meta-analysis. Pain Res Man :4276175. . https://doi.org/10.1155/2022/4276175, 2022.

47. Yanyun G, Lei H, Yi Z, Jing T, Kong W, Jingsong W: The effect of Pilates exercise training for scoliosis on improving spinal deformity and quality of life: Meta-analysis of randomized controlled trials. Medicine 100(39):e27254. https://doi.org/10.1097/MD.0000000000027254, 2021.

48. Bons S.C.S., Borg M.A.J.P., Van den Donk M. et al. NHG standard M54 Non-specific low back pain. Netherlands Huisarten Genootschap. Available at: https://richtlijnen.nhg.org/standaarden/aspecifieke-lagerugpijn Accessed July 2022.

49. Bussières AE, Stewart G, Al-Zoubi F, *et al.* The treatment of neck pain–associated disorders and whiplash-associated disorders: A clinical practice guideline. J Manipulative Physiol Ther 39(8):523-564, 2016. https://doi.org/10.1016/j.jmpt.2016.08.007

50. Finnish Medical Association Duodecim: Neck pain (adults). A working group set up by Duodecim of the Finnish Medical Society, Societas Medicinae Physicalis et Rehabilitationis Fenniae ry and the Finnish Society of General Medicine. Helsinki, Suomalainen Lääkäriseura Duodecim; 2017

51. Gauthier K, Dulong C, Argáez C: Multidisciplinary Treatment Programs for Patients With Chronic Non-Malignant Pain: A Review of Clinical Effectiveness, Cost-effectiveness, and Guidelines—An Update. Ottawa, CADTH; 2019

52. George SZ, Fritz JM, Silfies S, *et al.* Interventions for the management of acute and chronic low back pain: revision 2021: Clinical practice guidelines linked to the international classification of functioning, disability and health from the academy of orthopaedic physical therapy of the American Physical Therapy Association. J Orthop Sports Phys Ther 51(11):CPG1-CPG60, 2021. https://doi.org/10.2519/jospt.2021.0304

53. Kawakami M, Takeshita K, Inoue G, *et al.* Japanese Orthopaedic Association (JOA) clinical practice guidelines on the management of lumbar spinal stenosis, 2021—Secondary publication. J Orthop Sci 28(1):46-91, 2022. https://doi.org/10.1016/j.jos.2022.03.013

54. Kreiner DS, Matz P, Bono CM, *et al.* Guideline summary review: An evidence-based clinical guideline for the diagnosis and treatment of low back pain. Spine J 20(7):998-1024, 2020. https://doi.org/10.1016/j.spinee.2020.04.006

55. National Guideline Centre (UK): Low Back Pain and Sciatica in Over 16s: Assessment and Management. Manchester, National Institute for Health and Care Excellence; 2016. ⟨https://www.nice.org.uk/guidance/ng59⟩

56. Parfenov VA: Y. Chronic nonspecific (musculoskeletal) low back pain. Guidelines of the Russian Society for the Study of Pain (RSSP). Neurol Neuropsychiatr Psychosom 11(2S):7-16, 2019. https://doi.org/10.14412/2074-2711-2019-2S-7-16

57. Van Wambeke P, Desomer A, Ailliet L, *et al.* Low Back Pain and Radicular Pain: Assessment and Management. Good Clinical Practice (GCP). Brussels, Belgian Health Care Knowledge Centre (KCE); 2017

58. Avenell A, Stewart F, Grey A, *et al.* An investigation into the impact andimplications of published papers from retracted research: Systematic search of affected literature. BMJ Open 9:e031909, 2019.

59. Cochrane Pregnancy and Childbirth Review Group. Identifying and handling potentially untrustworthy trials in Pregnancy and Childbirth Cochrane Reviews. Available at: https://pregnancy.cochrane.org/news/identifying-and-handling-potentially-untrustworthy-trials-pregnancy-and-childbirth-cochrane Aceesed September 16, 2021.

60. Grey A, Bolland MJ, Avenell A, Klein AA, Gunsalus CK: Check for publication integrity before misconduct. Nature 577(7789):167-169, 2020. https://doi.org/10.1038/d41586-019-03959-6

61. Parker L, Boughton S, Lawrence R, Bero L: Experts identified warning signs of fraudulent research: A qualitative study to inform a screening tool. J Clin Epidemiol 151:1-17, 2022. https://doi.org/10.1016/j.jclinepi.2022.07.006

62. Weibel S, Popp M, Reis S, Skoetz N, Garner P, Sydenham E: Identifying and managing problematic trials: A research integrity assessment (RIA) tool for randomized controlled trials in evidence synthesis. Res Synth Methods 14(3):357-369, 2022. https://doi.org/10.1002/jrsm.1599

63. Williams ACdeC, Fisher E, Hearn L, Eccleston C: Evidence-based psychological interventions for adults with chronic pain: Precision, control, quality, and equipoise. Pain 160(8):2149-2153, 2021. https://doi.org/10.1097/j.pain.0000000000002273

64. O'Connell NE, Cook C, Wand BM, Ward SP: Clinical guidelines for low back pain. A critical review of consensus and inconsistencies across three major guidelines. Best Pract Res Clin Rheum 30(6):968-980, 2016. https://doi.org/10.1016/j.berh.2017.05.001

65. Hultcrantz M, Rind D, Akl EA, *et al.* The GRADE Working Group clarifies the construct of certainty of evidence. J Clin Epidemiol 87:4-13, 2017

66. Faraone SV: Interpreting estimates of treatment effects: Implications for managed care. P T 33(12):700-711, 2008