CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Psychotherapy for the Treatment of Acute Musculoskeletal Pain: A Review of Clinical Effectiveness and Guidelines

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	April 17, 2020
Report Length:	43 Pages

CADTH

Authors: Saadul Islam, Nina Frey

Cite As: Psychotherapy for the Treatment of Acute Musculoskeletal Pain: A Review of Clinical Effectiveness and Guidelines. Ottawa: CADTH; 2020 Apr. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

Abbreviations

ACT	Acceptance and commitment therapy
ACP	American College of Physicians
AE	Adverse event
AGREE II	Appraisal of Guidelines for Research & Evaluation II
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
BPI	Brief Pain Inventory
CAS	Cumulated Ambulation Score
CBT	Cognitive-behavioral techniques
CBPT	Cognitive behavior-based physical therapy
CRD	Centre for Reviews and Dissemination
DDS	Descriptor Differential Scale
EQ-5D	Euro Quality of life-5 Dimension
GPE	Global Perceived Effect
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
HTA	Health Technology Assessment
HRQoL	Health related quality of life
LBP	Low back pain
MCID	Minimal clinically important difference
MPQ	McGill Pain Questionnaire
NDI	Neck Disability Index
NSAIDS	Nonsteroidal anti-inflammatory drugs
NRS	Numeric Rating Scale
OA	Osteoarthritis
ODI	Oswestry Disability Index
PDI	Pain Disability Index
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QWB	Quality of Well-Being
RCT	Randomized controlled trial
RMDQ	Roland Morris Disability Questionnaire
SBU	Swedish Agency for Health Technology Assessment and Assessment of Social Services
SMD	Standardized mean difference
SIP	Sickness Impact Profile
SMR	Skeletal muscle relaxants
TKR	Total knee replacement
TUG	Timed up-and-go
VAS	Visual Analogue Scale
WAD	Whiplash-associated disorders
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

Context and Policy Issues

Musculoskeletal pain is characterized by pain commonly affecting the following areas – joints (e.g. arthritic conditions), bones (e.g. osteoporosis, fracture), muscle (e.g. sarcopenia), spine (back and neck pain), and multiple body areas or systems (e.g. lupus).[1] Musculoskeletal pain is generally considered chronic in nature if it persists for more than three months, whereas an accepted definition of acute or subacute pain is less common in the literature. However, short-term (≤ 3 months) pain can be categorized subacute if it lasts for seven to 12 weeks; therefore, pain shorter in duration (≤ 6 weeks) can be considered acute.[2]

Psychological therapies are aimed at influencing the psychosocial processes and are recommended alone or in combination with pharmacologic treatments for the management

of pain, disability or related symptoms.[3] Cognitive-behavioral techniques (CBT) is one of the most common type of psychotherapy, consisting of cognitive restructuring, reframing and reappraisal based on the individual needs of patients and their specific situation.[4] Acceptance and commitment therapy (ACT) is a psychological intervention that uses acceptance and mindfulness strategies, in combination with commitment and behavior change strategies, to increase psychological flexibility.[5] Relaxation techniques involve systematically instructing participants in progressive muscle relaxation, relaxing breathing techniques, hypnosis, music therapy, guided imagery, or autogenic training (desensitization-relaxation technique).[4] Finally, mindfulness techniques is characterized by the purposeful and nonjudgmental focus on the present moment, whereby one engages in awareness of bodily sensation, thoughts, or emotion; a practice that is thought to counteract negative emotional states. Examples of mindfulness include deep breathing, sitting meditation, yoga, and a body scan, in which attention is directed throughout different parts of the body.[6]

Psychotherapy for pain management is primarily targeted at improving physical, emotional, social, and occupational functioning rather than resolution of the underlying cause of pain or pain itself. These therapies have variable success in pain management resulting from the differences in their scope, duration, administration process, and goals.[3] The objective of the current report is to evaluate the evidence on the clinical effectiveness of psychological interventions versus other treatments for acute (lasting < 6 weeks) or subacute (lasting ≥ 6 weeks but ≤ 3 months) musculoskeletal pain. Additionally, evidence-based guidelines regarding the use of psychological therapies for the treatment of acute or subacute musculoskeletal pain will be reviewed.

Research Questions

1. What is the clinical effectiveness of psychological therapies for the treatment of individuals with acute or subacute musculoskeletal pain?
2. What are the evidence-based guidelines regarding the use of psychological therapies for the treatment of individuals with acute or subacute musculoskeletal pain?

Key Findings

A total of seven relevant publications were included in this report: five systematic reviews (two with meta-analysis), one cluster randomized controlled trial and one evidence-based guideline. Two of the five systematic reviews were aimed at postoperative pain; the remaining three were aimed at subacute low back pain (7–12 weeks), subacute neck pain (≤ 3 months), and all types of musculoskeletal pain. The randomized controlled trial was aimed at subacute low back pain (2–12 weeks), and the guideline provided recommendations on all forms of low back pain. Overall, cognitive-behavioral therapy and low back pain were the most studied psychological intervention and musculoskeletal condition, respectively.

Cognitive-behavioral therapy combined with physiotherapy appeared to provide functional improvements following back-surgery, without any impact on pain resolution. Psychotherapies based on relaxation or mindfulness techniques showed mixed results on pain following knee surgery; a firm conclusion could not be drawn due to widely variable intervention scheme.

With respect to musculoskeletal pain, psychotherapy combined with physiotherapy was shown to improve pain and disability resulting from musculoskeletal pain overall; however, these benefits were not found when low back pain, neck and whiplash-associated pain, and osteoarthritis-related pain were investigated separately. Cognitive-behavioral therapy was found to be beneficial in subacute neck pain, although the evidence was of low quality. There was some evidence that cognitive-behavioral therapies may reduce disability and improve body functions in patients with subacute low back pain, particularly when integrated with physiotherapy and personalized to patients' context; however, the effects on pain resolution was less pronounced. The clinical guideline made no reference to psychotherapies for the management of short-term low back pain, and instead recommended other forms of non-pharmacologic therapies since most patients achieve resolution naturally.

Overall, the included studies showed substantial heterogeneity in psychotherapies used and outcomes measures, making it difficult to compare findings across studies and to obtain an overall picture of the various psychotherapies for different types of musculoskeletal pain. Nevertheless, psychological therapies, most notably cognitive-behavioral therapy, has some clinical benefits in improving short-term pain and body functions resulting from surgery or musculoskeletal conditions when combined with other interventions aimed at improving body functions e.g. physiotherapy.

Methods

Literature Search Methods

A limited literature search was conducted by an information specialist on key resources including PubMed, the Cochrane Library, the University of York Centre for Reviews and Dissemination (CRD) databases, the websites of Canadian and major international health technology agencies, as well as a focused internet search. The search strategy was comprised of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords. The main search concepts were psychotherapy and acute pain. Search filters were applied to limit retrieval to health technology assessments, systematic reviews, meta-analyses, any type of clinical trial or observational study, and guidelines. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between Jan 1, 2015 and March 20th, 2020.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Patients (of any age) with acute (lasting < six weeks) or subacute (lasting ≥ six weeks but ≤ three months) musculoskeletal pain
Intervention	Psychotherapy (e.g., cognitive behavioural therapy, mindfulness-based cognitive therapy)
Comparator	Pharmacotherapy (e.g., ibuprofen, acetaminophen, opioids); non-pharmacological treatments (e.g., exercise, manual therapy); no treatment; wait list; any combination of the listed comparators (e.g., pharmacotherapy and exercise)
Outcomes	Q1: Clinical effectiveness (e.g., pain [e.g., Pain Numeric Rating Scale score), quality of life, incidence of subsequent chronic pain) Q2: Recommendations regarding best practices (e.g., treatment protocols, appropriate patient populations)
Study Designs	Health Technology Assessments, Systematic Reviews, Randomized Controlled Trials, Evidence-based Guidelines.

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2015. Therefore, the methodology could not be assessed for summary. One RCT by Rolving et al.[7] was excluded since this trial was captured by Nicholls et al.[5] and will therefore be discussed in the context of the systematic review. One systematic review by Chou et al.[8] met the inclusion criteria; however, it was not summarized separately as the study formed the evidence base of a clinical practice guideline. This guideline also included an updated search, therefore results of the systematic review were discussed in the context of the guideline.[9] Finally, guidelines with unclear methodology were excluded.

Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised by one reviewer using A MeaSurement Tool to Assess systematic Reviews (AMSTAR) II,[10] the clinical studies were critically appraised using the Downs and Black checklist,[11] and the guideline was critically appraised using the Appraisal of Guidelines for Research & Evaluation II (AGREE II) Checklist.[12] Summary scores were not calculated for the included studies; rather, the strengths and limitations of each included study were described narratively.

Summary of Evidence

Quantity of Research Available

A total of 381 citations were identified in the literature search. Following screening of titles and abstracts, 355 citations were excluded and 26 potentially relevant reports from the electronic search were retrieved for full-text review. In addition, five potentially relevant publications were retrieved from the grey literature search for full-text review. Of these 31 potentially relevant articles, 24 publications were excluded for various reasons, while seven publications met the inclusion criteria and were included in this report. These comprised of five systematic reviews, one RCT, and one clinical practice guideline. Appendix 1 presents the PRISMA[13] flowchart of the study selection." Additional references of potential interest are provided in Appendix 6.

Summary of Study Characteristics

A total of five relevant systematic reviews (including two with meta-analyses[14,15] and three with descriptive analysis),[2,5,16] one RCT,[17] and one evidence-based guideline[9] were identified for inclusion in this review. This report will be limited to evidence pertaining to the research questions listed above. Therefore, studies were only included if they assessed the clinical benefits of talk-based psychotherapies in patients where the short-term (≤ 3 months) musculoskeletal nature of the underlying pain was clearly discernible. Detailed study characteristics are available in Appendix 2, Table 2, and Table 3, and Table 4.

Study Design

The five included systematic reviews[2,5,14-16] had objectives and inclusion criteria that were wider in scope than the current report. The systematic reviews compared a variety of psychological therapies for the treatment of a number of surgical or disease-related musculoskeletal conditions. The systematic reviews were published between 2015 and 2019, and included randomized controlled trials (RCTs). Whale et al. (2019) [16] and Nicolls et al. (2018)[5] included 12 and five RCTs, respectively, and both investigated postoperative pain management using psychotherapies. The former was aimed at pain following total knee replacement (TKR) surgery, and the latter included all forms of post-surgical pain, of which four RCTs reporting pain from back-surgery were relevant. Mariano et al. (2018)[2] included six RCTs that investigated the clinical effectiveness of cognitive behavioural therapies (CBTs) in subacute low back pain (LBP). Monticone et al. (2015)[15] included 10 RCTs aimed at the effect of CBTs on subacute and chronic neck pain, of which two provided relevant information for this report. Silva Guerrero et al. (2018)[14] included findings from 34 RCTs on the effectiveness of psychotherapies on various musculoskeletal pain conditions.

The RCT by Mas et al.[17] was a single-blinded parallel-group longitudinal cluster trial designed to compare various pharmacologic and non-pharmacologic interventions in subacute non-specific LBP. Patients were randomized by cluster to minimize contamination since the intervention was delivered to groups, and the randomization unit was the primary healthcare centre. The physicians identified the patients for enrolment during consultation and were aware of the treatment allocation. Following this, patients were invited to participate if they met the eligibility criteria, without knowing their treatment allocation.

The single guideline included in this report was developed by the American College of Physicians (ACP)[9] that presented evidence and provided clinical recommendations on noninvasive pharmacologic and nonpharmacologic treatment for LBP. Intended for clinicians, the guideline was based on a systematic review[8] that included RCTs published up to February 2016. Authors of the guideline conducted an updated search through November 2016 in preparation of the recommendations.

Country of Origin

Authors of systematic reviews were based in UK,[16] USA,[2] Canada,[5] Australia,[14] and Italy.[15] The location of the primary studies was not provided in all systematic reviews.

The RCT by Mas et al.[17] was conducted in 39 primary healthcare centres in Barcelona, Spain.

The ACP guideline[9] was published by a group of scientists in the US.

Patient Population

All included systematic reviews consisted of adult (≥ 18 years of age) patients with a variety of postoperative or disease-related musculoskeletal and non-musculoskeletal pain. Whale et al.[16] included 12 RCTs on patients with short and long-term pain following TKR. Nicholls et al.[5] comprised of six trials including patients with post-surgical pain, of which five trials were relevant for this review that included patients who underwent back-surgery, including spinal fusion and laminectomy surgery, and experienced a number of pre-existing spinal pathologies e.g. back pain or lower-extremity pain (sciatica). Neither of these two systematic reviews defined an acute/subacute or chronic period of postoperative pain, therefore, evidence pertaining to three months of follow-up was summarized in this report. Mariano et al. included patients with subacute LBP, defined as LBP-related pain lasting 7–12 weeks, although studies included in the review had lower limits for subacute LBP of 3–8 weeks and upper limits of 10–12 weeks. Monticone et al. [15] included patients with both clinically diagnosed subacute (i.e. a documented history of pain lasting for at least one month and not longer than three months) or chronic (i.e. a documented history of pain lasting for at least three months) neck pain. Finally, the eligible trials in Silva Guerrero et al.[14] included diagnosed cases of an acute or chronic musculoskeletal pain condition (including osteoarthritis [OA], LBP, neck pain, whiplash-associated disorders [WAD], and temporomandibular joint syndrome).

The RCT by Mas et al.[17] included working age patients (aged 18–65 years) with subacute LBP (2–12 weeks) who did not have a history of LBP during the preceding six months.

The ACP guideline[9] provided treatment guidance for adult (aged ≥18 years) patients with acute (< 4 weeks), subacute (4 to 12 weeks), and chronic (> 12 weeks) low back pain (LBP), radicular low back pain, or symptomatic spinal stenosis, in both inpatient and outpatient settings.

Interventions and Comparators

Systematic reviews

The five included systematic reviews investigated the clinical benefits of a variety of psychological interventions regarding pain management. The duration and delivery of psychotherapy interventions varied substantially across studies, ranging from three to 26 sessions in total, from 30 minutes to 3 hours per session, delivered in person or by phone, in individual or group format, and were led by a nurse, physiotherapist, psychologist, or other health professionals. Additionally, psychotherapies to assess postoperative pain were administered before, during, or after surgery. Notwithstanding the above factors, the primary study details were not reported with adequate detail in some instances.

The systematic review by Whale el.[16] included a broad category of psychological interventions including CBT, behavioral interventions, ACT, social skills training, relaxation therapies, mindfulness, psychodynamic, counselling, and interpersonal therapies. The inclusion criteria of the systematic review by Nicolls et al.[5] consisted of all forms of psychological therapies; however, the included studies incorporated CBT-based psychotherapies with no studies identified for ACT or mindfulness.

Mariano et al.[2] and Monicone et al.[15] both investigated the effectiveness of CBT interventions, alone or in conjunction with other therapeutic modalities such as exercise or physical modalities. The CBT interventions encompassed a wide set of interventions that included cognitive reconditioning (e.g. cognitive restructuring, imagery, attention diversion,

relaxation techniques) and behavioral modifications of specific activities (e.g. operant treatment, pacing, graded exposure approaches) to modify and/or reduce the impact of pain and physical and psychosocial disability and to overcome barriers to physical and psychosocial recovery.

Silva Guerrero et al.[14] included all forms of psychological interventions in combination with physiotherapy to physiotherapy alone or usual care. Psychological interventions included in the primary studies involved cognitive- behavioral techniques, pain and stress management program, relaxation training, hypnosis, mindfulness or acceptance-based interventions, pain coping skills training, problem-solving (self-determination theory-based communication skills, solution finding approach), systematic desensitization, motivational interview, fear-avoidance and anxiety, work rehabilitation program, graded activity (behavioral-graded activity, workplace intervention graded activity, graded exposure, graded exercise), intensive group training, depression interventions, patient decision support package and multifaceted approach (education, self-management strategies and an active exercise component) – given alone or in combination with physiotherapy. Examples of usual care included self-care instructions, ultrasound, and education.

Randomized controlled trial

In the Mas et al. RCT,[17] patients in both the control and intervention group received guideline-based pharmacological treatment. Participants in the control group received usual clinical care, based on the Clinical Guidelines for Lumbar Spine Disorders in Adults published by the Catalan Institute of Health. Patients in the intervention group received usual care in addition to a psychological program (consisting of an educational booklet and some audio-visual materials). The intervention was administered by a physician and/or nurse, a psychologist and a physiotherapist – a program that lasted between 90 and 120-minute (total 10 hours). Following the completion of the study, the control group also received the educational booklet and the audio-visual material.

Clinical guideline

The ACP guideline[9] provided recommendations on noninvasive pharmacologic treatments (acetaminophen, nonsteroidal anti-inflammatory drugs [NSAIDs], opioids, skeletal muscle relaxants [SMRs], benzodiazepines, antidepressants, antiseizure medications, and systemic corticosteroids) and nonpharmacologic treatments (including psychological therapies, multidisciplinary rehabilitation, spinal manipulation, acupuncture, massage, exercise and related therapies, and various physical modalities) for low back pain. As such, topical pharmacologic therapies or epidural injection therapies were not included. The interventions were compared with each other or with placebo (drug trials), sham (functionally inert) treatments, or no treatment.

Outcomes

Systematic reviews

The systematic reviews included in the report measured outcomes using various questionnaires and scales. However, most of the studies did not provide any information on the description, validity, reliability, and minimal clinically important difference (MCID) of these questionnaires. This report will be limited to outcomes related to pain, disability, and quality of life in accordance with the inclusion criteria. As such, psychological outcomes and factors (e.g. self-efficacy, catastrophizing, fear of movement/pain, kinesiophobia, anxiety, depression) as well as outcomes related to work absenteeism (e.g. return to work, days of

sick leave) and healthcare utilization (days of hospitalization) will not be reported here. The outcomes measured in the systematic reviews can be broadly categorized as follows:

Pain related outcomes: All systematic reviews measured pain-related outcomes as a primary or secondary outcome, using a number of questionnaires. The questionnaires used in the systematic reviews included: Pain Visual Analogue Scale (VAS, range 0–10), Numeric Rating Scale (NRS, range 0 i.e. no pain to 10 i.e. maximum pain), The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) Pain Scale, Short-form McGill Pain questionnaire, Brief Pain Inventory (BPI), Short Form 36 – Body Pain Scale, Low Back Pain Rating Scale, resolution of pain (using Descriptor Differential Scale [DDS]) and modified von Korff scale. In addition, an indirect estimation of pain relief was reported in some reviews by use of rescue analgesics. Nicholls et al.[5] reported that the pain and physical functioning outcomes measured in the review were in line with the IMMPACT core outcome measures for chronic pain. Additionally, an absolute decrease in ≥ 35 points in VAS and a 2-point reduction on the NRS was reported as an MCID for these scales.

Functional and disability-related outcomes: A number of outcomes reported in the included systematic reviews were aimed at measuring functionality and disability. The reviews captured trials that reported outcomes using a combination of the following questionnaires or scales: disability using the Oswestry Disability Index (ODI), Roland Morris Disability Questionnaire (RMDQ), Pain Disability Index (PDI), Neck Disability Index (NDI, range 0 i.e. no disability to 50 maximal disability), Neck Pain and Disability Scale, Quebec Back Pain Disability Scale; ambulation using the Cumulated Ambulation Score (CAS); and mobility using the five-chair-stand test, timed up-and-go (TUG) test, and 10-minute walk test.

Health-related Quality of Life and psychological well-being/status: A number of generic and disease-specific health related quality of life (HRQoL) indicators were used in the systematic reviews, including: Short Form version 12 or 36 (SF-12 and 36), life satisfaction (Cantril's Ladder Scale), Euro Quality of life-5 Dimension (EQ-5D), health status measure by Quality of Well-Being (QWB), and Sickness Impact Profile (SIP).

In addition, the following global measures of health indicators were used: Global improvement or perceived recovery (overall improvement, proportion of participants recovered, subjective improvement of symptoms), satisfaction with treatment (e.g. Global Perceived Effect (GPE), reduction in frequency or number of medications used.

Randomized controlled trials

The RCT by Mas et al.[17] measured disability (using a translated and validated version of the RMDQ) as the primary outcome. This is a 24-item scale, with lower score indicating less disability. The authors reported a difference of at least 2.5 points in RMDQ score compared to the baseline value to be clinically meaningful for subacute and chronic pain. In addition, the Spanish version of the McGill Pain Questionnaire (MGPQ) was used to measure pain (as a secondary outcome), which assesses the following three parameters with three dimensions (sensorial, affective and evaluative): Total Intensity Score (scale 0–14), Current Intensity Score (scale 0–5) and Visual Analogical Scale (VAS, scale 0–10). The mental and physical HRQoL was measured using the Spanish version of SF-12 (scale 0–100; lower scores indicate worse HRQoL).

Clinical guideline

The ACP guideline[9] and supporting SR[8] included the following outcomes – reduction or elimination of LBP (measured using VAS), improvement in back-specific and overall function (measured by ODI, RMDQ), improvement in HRQoL, reduction in work disability, global improvement, number of back pain episodes or time between episodes, patient satisfaction, and adverse effects. Timing of outcomes was stratified as long-term (≥ 1 year) and short-term (≤ 6 months).

Summary of Critical Appraisal

Additional details regarding the strengths and limitations of included publications are provided in Table 5, Table 6, and Table 7 of Appendix 3.

Systematic reviews

The five included systematic reviews were conducted and reported in accordance with established guidelines for systematic reviews (PRISMA statement and Cochrane guidance). All five reports had a clearly defined objective and rationale, inclusion and exclusion criteria, and included flow charts illustrating study selection and provided reasons for study exclusion. Key search terms and search strategies were provided in all reviews, increasing their reproducibility. A comprehensive and thorough literature review was conducted, covering multiple databases, with little or no restrictions placed on publication date. Study selection, data extraction, and study quality assessment were well-documented and generally done in duplicate for all but one report; Mariano et al.,[2] where data extraction was not done in duplicate and risk of bias assessment was not done at all. In the four other systematic reviews, established tools from the Cochrane research group was used to assess the risk of bias and quality of evidence (Cochrane risk of bias tool and the GRADE approach). Finally, the authors of all five systematic reviews disclosed their sources of funding; with no noteworthy conflicts of interest.

Description of important characteristics of the included studies was adequately provided. All systematic reviews were generally conducted with high scientific rigour; however, the quality of included studies within each systematic review varied. Overall, the risk of bias of the included RCTs in each systematic review was mixed, as reported by the respective authors, and most primary studies had multiple domains with low, moderate, high or unclear risk of bias. In general, the domains associated with blinding of the participants and outcome assessors (i.e. performance and detection bias) were rated as high in risk in most of the included trials. This can be expected owing to the nature of receiving psychological interventions which poses a methodological challenge in implementing and maintaining blinding. Nonetheless, the quality of the studies and evidence was adequately considered in interpreting results and formulating conclusions.

The statistical methods for pooling results for pairwise comparison were appropriately done in Silva Guerrero et al.[14] and Monticone et al.[15] The remaining three systematic reviews reported a high degree of heterogeneity in the conduct of the trials, rendering it impossible to pool data, and presented results in a narrative manner. Both meta-analyses assessed statistical heterogeneity using the I^2 statistic. In Monticone et al.,[15] a random-effects model was used when $25\% < I^2 < 75\%$ indicating moderate heterogeneity, otherwise a fixed-effects model was used. Silva Guerrero et al.[14] inspected between-study heterogeneity for each analysis, which was taken into account when assessing the quality of evidence; however, the decision whether or not to perform meta-analysis was not dependent on I^2 value. Assessments of publication bias were planned in Monticone et al.[15] and Silva

Guerrero et al.[14] using Funnel plot (visually) and Egger test and Harbord test (quantitatively). However, it could not be assessed in Monticone et al.[15] due to the small number of included studies (< 10); whereas Silva Guerrero et al.[14] found evidence of plot asymmetry for a number of subgroup analyses which indicated the possibility of small study bias.

Randomized controlled trials

The RCT by Mas et al.[17] had a clearly defined study objective, eligibility criteria, methods for patient recruitment, randomization scheme and choice of interventions. In addition, the main outcomes, potential confounders, baseline patient characteristics, and main findings were also clearly described, increasing the strength of reporting. The trial was single-blinded, with treatment allocation masked from the patients and analysts but not from the study investigators. However, the assessment of study outcomes as well as the diagnosis of patients for eligibility appeared to have been done in an objective manner. Statistical tests were conducted in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines; and analyses were adjusted for significant confounders and significant interaction variables. Study participants, care providers or investigators, and health care settings appeared to be representative of the population and care settings of interest, increasing the external validity of the studies.

As for methodological limitations, the RCT recruited a smaller number of patients than was estimated to determine a clinically meaningful effect with adequate power. There were no other noteworthy limitations.

Clinical guidelines

The ACP guideline[9] had a clearly described scope and purpose through specific descriptions of the overall objectives, health questions, intended users, and target populations. The views and preferences of the target population were sought; as the guideline and the accompanying evidence reviews underwent internal peer review and external stakeholder feedback prior to publication. The guideline was developed using rigorous systematic methodology and was based on a systematically reviewed and critically appraised body of clinical evidence. Recommendations in the guideline were clear and unambiguous, and were accompanied by a grading of the associated evidence and a measure of strength of the recommendation. Details regarding the exact methods used to form the recommendations and information regarding updating the guideline were provided. However, facilitators and barriers to its application, potential resource implications, implementation guidance, and monitoring or auditing criteria were not described or unclear in the guideline. Conflicts of interest were addressed.

Summary of Findings

Results are presented for studies (including primary studies in systematic reviews) that reported on the use of psychological interventions for the treatment of acute/subacute musculoskeletal pain. As such, results of meta-analysis are described narratively for the primary studies where acute/subacute pain were analyzed separately from chronic pain, and where musculoskeletal pain was distinguished from non-musculoskeletal pain. Likewise, results of the systematic reviews were limited to primary trial(s) where the psychotherapy, musculoskeletal condition and duration of pain were clearly identified to be in line with the objective of this report. Given the wide variation in trial design, eligibility criteria, methods, study length, sample size, nature and duration of psychotherapies, and outcome variables, no statistical meta-analysis was performed for Whale et al., Nicholls et

al., and Mariano et al. Appendix 4 presents a table of the main study findings and authors' conclusions.

Clinical Effectiveness of psychological therapies for the treatment of individuals with acute or subacute musculoskeletal pain

Results of the included studies are summarized by the nature of musculoskeletal pain, i.e. postsurgical pain, and musculoskeletal pain by body part or disease type. Recognizing the different types of psychotherapies used in the primary studies included in the systematic reviews, and the heterogeneity in terms of how psychotherapies were performed and administered, a short description of the psychotherapies used in each instance are given in Appendix 4.

Psychological therapies for the treatment of post-surgical pain

CBT programs

Whale et al.[16] reported two RCTs that evaluated the effectiveness of CBT-based programs, of which one reported results at six months of follow-up only (therefore 3-month data were not available). The other RCT evaluated a CBT-based program; although the study aimed to assess anxiety and depression compared to standard care, it also assessed pain using the WOMAC pain score. This RCT found no between-group differences found at four months of follow-up in pain measured using the WOMAC pain score.

Nicholls et al.[5] reported five relevant RCTs assessing psychotherapies on pain and function-related outcomes among back-surgery patients; two used CBT integrated with physiotherapy. These trials included a psychomotor therapy and a cognitive behavior-based physical therapy (CBPT), both of which had physiotherapy regimens integrated into behavioral self-management, cognitive restructuring, and relaxation training. In both trials, the CBT-integrated physiotherapy intervention was associated with a statistically significant improvement in pain outcomes at three months from baseline and compared to standard of care (education and exercise support), as measured by VAS for back pain and BPI subscales. However, the change from baseline at three months were not within the reported MCID for either scale (absolute decrease of ≥ 35 in VAS and a 2-point reduction on the NRS). The CBT-integrated physiotherapy intervention was also associated with a statistically significant improvement in functional outcomes at three months from baseline and compared to standard of care, as measured by ODI for disability and performance-based mobility outcomes (five-chair-stand test, TUG test, and 10-minute walk test).

Nicholls et al.[5] also identified three RCTs that evaluated the impact of CBT alone compared to education and/or exercise support on pain and function-related outcomes among back-surgery patients. One trial reported a statistically significant improvement associated with CBT compared with exercise, as measured by the NRS for back and leg pain and the SF-36-BP for body pain. However, two studies reported no statistically significant improvements in pain outcomes associated with group CBT over an acute (7 days post-surgery) or subacute (3 months) follow-up period, as measured by the PRS, NRS and through consumption of rescue analgesics. In terms of functional outcomes, all three trials reported a statistically significant improvement in disability associated with CBT intervention compared to exercise, as measured by ODI and mobility (including timed tests for walking, rising and sitting from a chair, and getting in and out of bed).

Music therapy

Whale et al.[16] reported five RCTs that evaluated the effectiveness of music therapy for reducing acute postoperative pain following TKR. Four trials that used a number of 5–30 minute preselected music therapy, delivered by music-therapists or others, at various timepoints before, during, and after surgery showed no difference in pain severity compared to no music therapy. One of the five RCTs reported a statistically significant lower mean VAS pain scores through 24 hours post-surgery compared patient-selected music during surgery to white noise.

Guided imagery and music

Whale et al.[16] reported one RCT that evaluated the effectiveness of 20-minute guided imagery and commercially available recordings on outcomes post-surgery; however, no comparative findings were reported between the intervention groups at various timepoints up to six months post-surgery.

Hypnosis

Whale et al.[16] reported a single RCT that evaluated the effectiveness of a 35-minute pre-recorded hypnotic audio recording, usual care, and minimal treatment (psychoeducation, diaphragmatic breathing, relaxing music) on pain outcomes 24h post-surgery, and found small difference between treatment groups in mean NRS ratings.

Progressive muscle relaxation with biofeedback

Whale et al.[16] reported one RCT that evaluated the effectiveness of biofeedback-assisted progressive muscle relaxation skills on pain before TKR surgery and during continuous passive motion therapy post-surgery compared to standard continuous passive motion therapy, and showed a statistically significantly lower NRS score in the intervention group.

Pain coping skills program

Whale et al.[16] reported one RCT comparing the effectiveness of a pain coping skills program (which included sessions on cognitive restructuring, thought identification and challenging, self-calming and relaxation techniques, and activity management) to arthritis education and to usual care found no differences in mean WOMAC pain treatment scores up to 12 months post-surgery.

Motivational interviewing

Whale et al.[16] reported one RCT that compared the effectiveness of an enhanced postoperative recovery program involving motivational interview with physiotherapy, and found no differences in mean WOMAC pain treatment scores up to six months post-surgery.

Psychological therapies for the treatment of musculoskeletal pain of different body parts

Silva Guerrero et al.[14] conducted a meta-analysis comparing the treatment effects of psychological interventions combined with physiotherapy, physiotherapy alone or usual care by combining data from 34 studies, of which one study included participants with acute (< 6 weeks) musculoskeletal pain, two studies included participants with subacute (6 to 12 weeks) pain, and 23 studies included a mix of participants with acute/subacute and chronic musculoskeletal pain. Furthermore, results were provided for all types of musculoskeletal pain combined as well as by different body parts (provided below).

General/non-specific musculoskeletal pain

Overall, results showed physiotherapists-delivered psychological interventions in combination with physiotherapy to be a statistically significant better alternative in the management of short-term pain (26 studies) and disability outcomes (29 studies) compared with physiotherapy alone or usual care. Based on GRADE quality of evidence, the evidence was of moderate quality for pain outcomes and low quality for disability outcomes. The effect size was considered small for both types of outcomes by the authors. Notably, results were presented in an aggregate manner for all psychological interventions, and for any period from four to 16 weeks.[14]

Low Back Pain

Silva Guerrero et al.[14] reported no statistically significant difference between the three intervention arms (physiotherapists delivered psychological interventions and physiotherapy/usual care interventions) for pain (15 studies) or disability (17 studies) in patients with LBP at short-term follow-up.

Mariano et al.[2] identified six RCTs that investigated the effectiveness of a number of CBT-based interventions in patients with subacute LBP (7–12 weeks), measured by improvements in pain and disability outcomes. However, the results were not reported and discussed with adequate detail. Three RCTs exclusively included patients with subacute LBP, all of which favored individualized CBT over control treatment with respect to improving disability outcomes (RMDQ), reducing pain intensity (VAS), and greater resolution of pain and restoration of function (DDS and SIP, respectively). One RCT provided data for subacute and chronic LBP separately; however, the outcomes were not pain related, therefore not reported here. Two other RCTs pooled subacute and chronic LBP cases and showed inconsistent findings – one trial reported a statistically significant improvement in disability (RMDQ), pain (von Korff score), and HRQoL (SF-12) among patients receiving CBT, whereas the other RCT did not find the comparative benefits of CBT to be statistically significant.

The cluster RCT by Mas et al.[17] reported that patients with non-specific subacute LBP who received the biopsychosocial multidisciplinary intervention (physiotherapy, CBT and medication) showed a statistically significant improvement in disability metric (measured by RMDQ score) compared to usual clinical care at three and 12 months of follow-up. The MCID defined by the authors (2.5 points from baseline level) was reported in the intervention group at both timepoints, whereas the control group achieved the MCID in the RMDQ score only at month 12. With respect to the level of pain, the intervention group showed a marginal difference in pain intensity (measured by MPQ) at 12 months, current intensity score and VAS score at three months. The two groups were not significantly difference with respect to HRQoL, as measured by SF-12.

Neck pain or whiplash-associated disorders

Silva Guerrero et al.[14] reported no statistically significant difference between the three intervention arms (physiotherapy delivered with psychological interventions, physiotherapy alone, and usual care) for pain (7 studies) or disability (9 studies) in patients with WAD and neck pain at short-term follow-up.

Monticone et al.[15] identified two RCTs that assessed the effectiveness of CBT on subacute neck pain, which were subsequently combined in a meta-analysis. Results showed that CBT was better than other interventions (information booklet, didactic discussion, manual

therapy) for improving pain, although the evidence was of low quality (i.e. “Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate”).[15] Additionally, no difference was found in terms of disability based on low-quality evidence.

Osteoarthritis

Silva Guerrero et al.[14] reported no statistically significant difference between the three intervention arms (physiotherapists delivered psychological interventions and physiotherapy/usual care interventions) for pain (4 studies) or disability (3 studies) in patients with OA at short-term follow-up.

Guidelines regarding the use of psychological therapies for the treatment of individuals with acute or subacute musculoskeletal pain

No evidence or recommendation was provided regarding the use of psychological interventions for the treatment of acute or subacute LBP.[9] The authors reported that there was insufficient evidence to determine the effectiveness of psychological therapies in acute or subacute LBP.

Limitations

Overall, the studies included in this report were generally conducted with robust methodology, however, there are some limitations. The systematic reviews included primary studies that varied significantly with respect to study conduct, in particular the way interventions were administered and outcomes were measured. The variability and lack of standardization in psychological therapies was apparent from the different frequency, duration, timing, delivery personnel, and combination of therapies the RCTs used in each systematic review. Additionally, the RCTs used a variety of scales and questionnaires, with unclear or unreported validity, reliability, and MCID. With the exception of some commonly reported questionnaires to assess outcomes for pain and disability (e.g. VAS, NRS, WOMAC, ODI); the applicability and interpretability of the remaining outcomes are unclear. Given the heterogeneity in psychotherapies and outcomes across trials, it is difficult to separate the effects of a psychological intervention from its method of administration when trials of different nature are discussed together in the context of a systematic review. This also makes it difficult to obtain an overall picture of the various psychotherapies for different musculoskeletal pain.

Of the five systematic reviews, Whale et al.[16] and Nicholls et al.[5] assessed short and long-term pain resulting from surgery (TKR and back surgery, respectively). The scope of this report limited the follow-up period up to three months; however, some of the primary studies in these systematic reviews reported longer-term follow-up. There is a possibility that the effects seen during the 3-month timepoint may not be replicated at a longer timepoint, presenting additional challenges in extrapolating the short-term findings beyond three months.

Of the three systematic reviews that assessed short-term pain resulting from a musculoskeletal condition, Silva Guerrero et al.[14] combined all forms of psychological therapies into one intervention, therefore the effects of different psychotherapies could not be ascertained separately. Additionally, the studies did not report adequate details on the interventions and results in some instances.

Conclusions and Implications for Decision or Policy Making

A total of seven relevant publications were included, which comprised five systematic reviews,[2,5,16] two with a meta-analysis,[14,15] one single-blind cluster RCT,[17] and one evidence-based guideline.[9] Two of the five systematic reviews were aimed at postoperative pain (3 months of data presented); the remaining three were aimed at subacute LBP (7–12 weeks), subacute neck pain (≤ 3 months), and all types of musculoskeletal pain. The RCT was aimed at subacute LBP (2–12 weeks), and the guideline provided recommendations on all forms of LBP.

Findings from a systematic review[5] showed that compared to education and exercise support, CBT had a beneficial effect on disability and mobility for short-term postoperative pain (from back-surgery) when combined with physiotherapy. However, physiotherapy-integrated CBT did not show a clinically meaningful difference in pain-related outcomes. Another systematic review[16] found no effect of CBT alone on pain following knee surgery. Among other psychotherapies, guided imagery and music, hypnosis, pain coping skills program, and motivational interviewing showed no effect on short-term pain following knee surgery, whereas music therapy (only when selected by patients) and progressive muscle relaxation with biofeedback showed some benefits.[16]

With respect to musculoskeletal pain, one systematic review[14] with the most comprehensive set of studies reported that psychological interventions combined with physiotherapy has small but noticeable improvement in pain and associated disability in the short-term compared to physiotherapy alone or standard care, based on moderate and low-quality evidence, respectively. The same systematic review also reported no comparative benefits of physiotherapy-combined psychotherapy in short-term pain resulting from LBP, neck pain and WAD, as well as OA.[14]

One systematic review[2] described the effectiveness of CBT-based interventions in patients with subacute LBP. In general, CBT-based interventions were associated with improvements in pain, disability, function, and possibly in quality of life, particularly when individualized to patients' needs and context.[2] The only RCT[17] included in the report showed that among patients with subacute LBP, the combined intervention of physiotherapy, CBT, and medication resulted in a clinically meaningful improvement in disability, but had marginal benefits with respect to pain, and no benefits on quality of life. The clinical guideline[9] made no reference to psychotherapies for the management of acute/subacute LBP, and instead indicated the use of other non-pharmacologic therapies given most patients with acute/subacute LBP achieve resolution naturally. In terms of other short-term musculoskeletal pain, one well-conducted systematic review[15] reported CBT to be a better treatment option compared to other interventions for improving subacute neck pain.

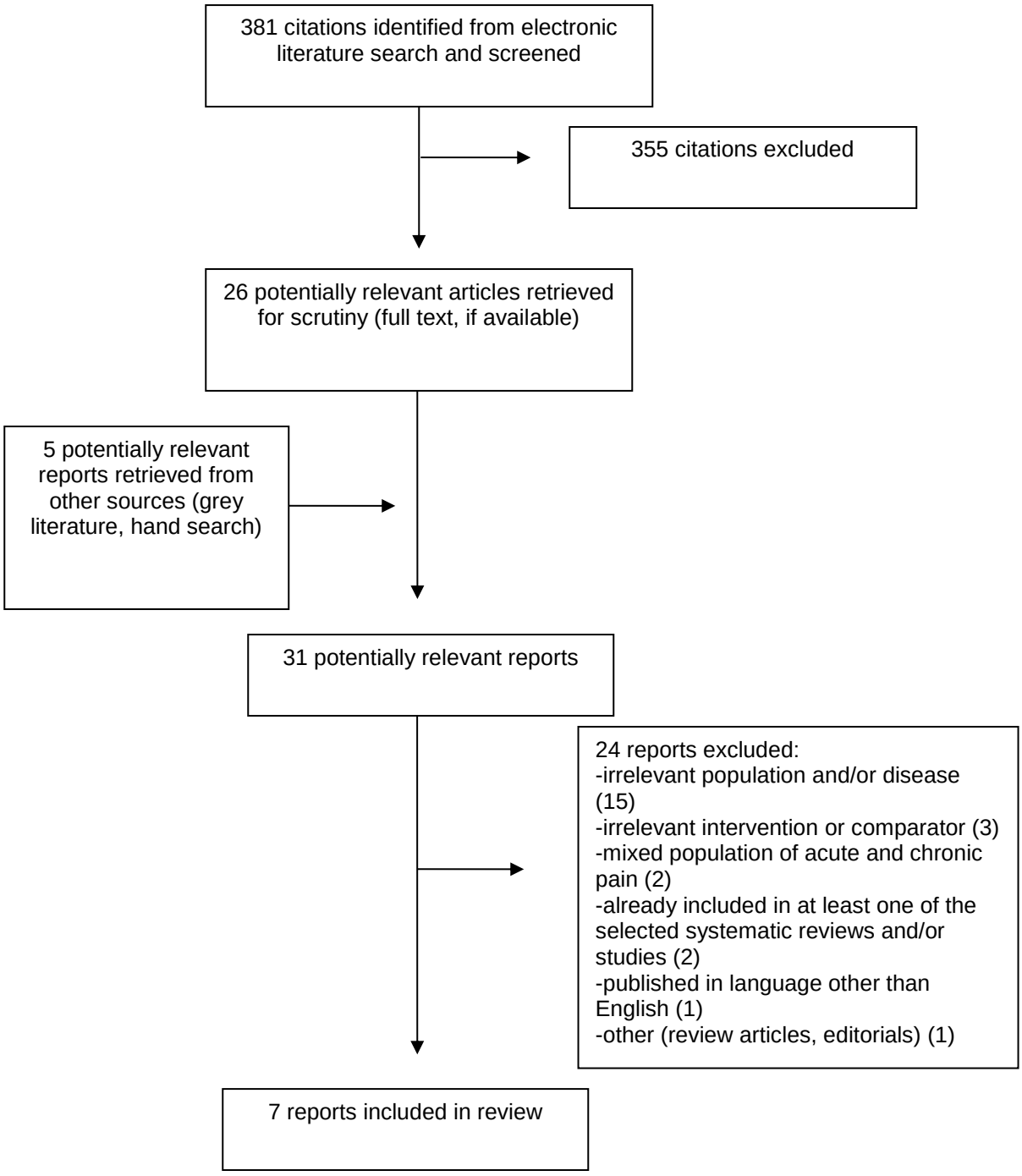
Overall, the included studies were of high quality, generally conducted with robust, well-reported methodology, although the primary studies comprising the evidence base of the systematic reviews and clinical guideline were heterogeneous in nature with respect to study conduct, interventions, and outcomes. The psychological therapies showed variability and lack of standardization with respect to the frequency, duration, timing, delivery personnel, and use of other therapies in combination with psychotherapies. The outcomes were measured with a variety of scales/questionnaires, the clinical use, validity and MCID of some were unclear or unreported. Given the heterogeneity in psychotherapies and outcomes, it is difficult to compare findings across studies and to obtain an overall picture of the various psychotherapies for different musculoskeletal pain. Nevertheless, psychological

therapies, most notably cognitive-behavioral therapy, combined with other interventions, such as physiotherapy, has some clinical benefits in improving short-term pain and body functions resulting from surgery or musculoskeletal conditions.

References

1. Musculoskeletal conditions. Geneva (CH): World Health Organization; 2019 Nov: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions. Accessed 2020 Apr 17.
2. Mariano TY, Urman RD, Hutchison CA, Jamison RN, Edwards RR. Cognitive Behavioral Therapy (CBT) for Subacute Low Back Pain: a Systematic Review. *Current pain and headache reports.* 2018;22(3):15.
3. Markozannes G, Aretouli E, Rintou E, et al. An umbrella review of the literature on the effectiveness of psychological interventions for pain reduction. *BMC Psychol.* 2017;5(1):31.
4. Ziehm S, Rosendahl J, Barth J, Strauss BM, Mehnert A, Koranyi S. Psychological interventions for acute pain after open heart surgery. *The Cochrane database of systematic reviews.* 2017;7:CD009984.
5. 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.* 2018;9:49-64.
6. Miller-Matero LR, Coleman JP, Smith-Mason CE, Moore DA, Marszalek D, Ahmedani BK. A Brief Mindfulness Intervention for Medically Hospitalized Patients with Acute Pain: A Pilot Randomized Clinical Trial. *Pain medicine (Malden, Mass).* 2019;20(11):2149-2154.
7. Rolving N, Nielsen CV, Christensen FB, Holm R, Bunger CE, Oestergaard LG. Preoperative cognitive-behavioural intervention improves in-hospital mobilisation and analgesic use for lumbar spinal fusion patients. *BMC Musculoskelet Disord.* 2016;17:217.
8. Chou R, Deyo R, Friedly J, et al. Nonpharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med.* 2017;166(7):493-505.
9. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of P. Noninvasive Treatments for Acute, Subacute, and Chronic Low Back Pain: A Clinical Practice Guideline From the American College of Physicians. *Ann Intern Med.* 2017;166(7):514-530.
10. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008. http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf. Accessed 2020 Apr 17.
11. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf. Accessed 2020 Apr 17.
12. Higgins JPT, Green S, editors. Figure 15.5.a: Drummond checklist (Drummond 1996). *Cochrane handbook for systematic reviews of interventions*. London (GB): The Cochrane Collaboration; 2011: http://handbook-5-1.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm. Accessed 2020 Apr 17.
13. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.
14. Silva Guerrero AV, 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. *The Clinical journal of pain.* 2018;34(9):838-857.
15. Monticone M, Cedraschi C, Ambrosini E, et al. Cognitive-behavioural treatment for subacute and chronic neck pain. *The Cochrane database of systematic reviews.* 2015(5):CD010664.
16. Whale K, Wylde V, Beswick A, Rathbone J, Vedhara K, Gooberman-Hill R. Effectiveness and reporting standards of psychological interventions for improving short-term and long-term pain outcomes after total knee replacement: a systematic review. *BMJ Open.* 2019;9(12):e029742.
17. Mas RR, Lopez-Jimenez T, Pujol-Ribera E, et al. Effectiveness of a multidisciplinary BIOPSYCHOSOCIAL intervention for non-specific SUBACUTE low back pain in a working population: a cluster randomized clinical trial. *BMC Health Serv Res.* 2019;19(1):962.
18. Agree Next Steps Consortium. The AGREE II Instrument. [Hamilton, ON]: AGREE Enterprise; 2017: https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf. Accessed 2020 Apr 17.
19. Garland EL, Baker AK, Larsen P, et al. Randomized Controlled Trial of Brief Mindfulness Training and Hypnotic Suggestion for Acute Pain Relief in the Hospital Setting. *J Gen Intern Med.* 2017;32(10):1106-1113.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Reviews and Meta-Analyses

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
Whale et al. 2019 UK[16]	12 RCTs published up to 9 May 2019	Adults undergoing primary TKR (N = 1299, range 24 – 402 participants)	**Interventions:** Relaxation/mindfulness (music therapy, hypnosis, progressive muscle relaxation with biofeedback), multimodal therapy of guided imagery and music, cognitive/behavioral (CBT-based programs, and a postoperative management program comprising motivational interviewing to improve self-efficacy and goal attainment), and a combination of relaxation/mindfulness and cognitive (pain coping skills program) intervention **Comparator:** active treatment or control treatment (e.g., standard care, placebo, no treatment).	**Clinical outcomes:** Postoperative knee pain severity after TKR, measured using Pain VAS, NRS, WOMAC Pain Scale, and the Short-form MPQ(primary) Function, HRQoL, and psychological well-being/status (secondary) SAE **Follow-up:** no time limit placed on assessment duration/ follow-up
Mariano et al. 2018 USA[2]	6 RCTs published within the past 20 years up to October 2017.	Patients with subacute LBP (7–12 weeks)	**Intervention:** CBT (Individual education, coping strategies, reassurance, individual biopsychosocial treatment, relaxation, individual psychotherapy, individual, custom program including problem solving, coping skills, relaxation, group therapy addressing "behaviors and beliefs about physical activity and avoidance of activity", "cognitive educational program") 2–6 sessions	**Clinical outcomes:** Pain: VAS, modified von Korff scale Disability: RMDQ, "Recovery" defined as resolution of pain (DDS) and restoration of function (SIP), HRQoL: Health status (QWB), EQ-5D, SF-12, SF-36, life satisfaction (CLS) **Follow-up:** 18 weeks – 24 months

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
Nicholls et al. 2018 Canada[5]	6 RCTs published up to February 2017, 4 were relevant and reported back surgery	Adult (aged ≥ 18 years) back-surgery patients (5 studies, 4 included spinal fusion, 1 with laminectomy surgery) and cardiac-surgery patients with comorbid major or minor depression (1 study) (sample size range 53–130)	**Intervention:** Cognitive therapy, mindfulness, and ACT (CBT and CBT-integrated with physical therapy) delivered pre, peri, or postoperatively. Duration ranged from 3 – 8 sessions, 30 minutes – 3 hours/session, and delivered in person or by phone, in individual or group format, and were led by a nurse, physiotherapist, or psychologist.	**Clinical outcomes:** Pain: Pain interference and/or severity measured by VAS, NRS, BPI, SF-36 BP scale, and LBP-RS, and consumption of rescue analgesics Functional outcomes: disability measured by ODI, ambulation using CAS, and mobility using the five-chair-stand test, timed up-and-go test, and 10 min walk test **Follow-up:** First postoperative week, 2 months to 2–3 years
Silva Guerrero et al. 2018 Australia[14]	34 RCTs identified through May 2016, of which 30 contributed to meta-analysis, 1 study included acute (<6 wk) pain only, 2 studies included participants with subacute (6 to 12 wk) pain only, 23 studies included participants with musculoskeletal pain of varying duration.	Adult patients (≥ 18 years) with a diagnosis of an acute or chronic musculoskeletal pain condition (OA, LBP, neck pain or WAD, temporomandibular joint syndrome) (N = 4936)	**Intervention:** Physiotherapist delivered psychological interventions (CBT, stress management, relaxation training, hypnosis, mindfulness or acceptance-based interventions, coping skills training, problem-solving, systematic desensitization, motivational interview, and anxiety and depression interventions) combined with physiotherapy **Comparator:** physiotherapy alone or usual care or no treatment	**Clinical outcomes:** Pain: VAS, NRS, modified von Korff scale Disability: RMDQ, ODQ/ODI, PDI, NDI, WOMAC, QBPDS **Follow-up:** short-term (4 to 16 weeks), long-term (26 to 52 weeks).
Monticone et al. 2015 Italy[15]	10 RCTs published up to November 2014, 2 studies included patients with subacute neck pain	Adults with subacute (pain lasting for ≥ 1 month and ≤ 3 months) and chronic neck pain (pain lasting for ≥ 3 months) (N = 836 total, n = 337 participants with subacute neck pain)	**Intervention:** CBT with or without another intervention (e.g. physiotherapy) **Comparator:** Placebo, no treatment (information booklet), or waiting list controls, and other types	**Clinical outcomes:** Pain: NRS Disability: NDI **Follow-up:** Short-term follow-up

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
			of interventions (didactic discussion, manual therapy)	

ACT = Acceptance and commitment therapy; BPI = Brief Pain Inventory; CAS = Cumulated Ambulation Score; CBT = Cognitive-behavioral techniques; CLS = Cantril's Ladder Scale; DDS = Descriptor Differential Scale; EQ-5D = Euro Quality of life-5 Dimension; HRQoL = Health related quality of life; LBP = low back pain; MPQ = McGill Pain Questionnaire; NDI = Neck Disability Index; NRS = Numeric Rating Scale; OA = Osteoarthritis; ODI = Oswestry Disability Index; RCT = randomized controlled trial; TKR = total knee replacement; QWB = Quality of Well-Being; RMDQ = Roland Morris Disability Questionnaire; SAE = serious adverse events; SIP = Sickness Impact Profile; SF-12/36 = Short Form 12/36; VAS = visual analogue scale; WAD = Whiplash-associated disorders; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Table 3: Characteristics of Included Primary Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention and Comparator(s)	Clinical Outcomes, Length of Follow-Up
Mas et al. 2019[17] Spain	Analyst-blinded longitudinal multicenter cluster RCT	Working population of patients (aged 18–65 years, mean age 46.8 years) with subacute (2–12 weeks), non-specific LBP (N = 501; control group = 239, intervention group = 262)	**Intervention:** Usual clinical care + a biopsychosocial multidisciplinary intervention, consisting of physiotherapy, CBT and medication **Comparator:** Guideline-based pharmacological treatment	**Clinical outcomes:** Disability: RMDQ (primary) Pain intensity: MPQ and Quality of Life: SF-12 (secondary) **Follow-up:** 12 months

CBT = cognitive-behavioural therapy; MGPQ = McGill Pain Questionnaire; LBP = low back painRCT = randomized controlled trial; RMDQ = Roland Morris Disability Questionnaire; SF-12 = Short Form 12-Item

Table 4: Characteristics of Included Guidelines

Intended Users, Target Population	Intervention and Practice Considered	Major Outcomes Considered	Evidence Collection, Selection, and Synthesis	Evidence Quality Assessment	Recommendations Development and Evaluation	Guideline Validation
American College of Physicians, 2017[9]						
Intended users: all clinicians Target population: adults (≥18 years) with	noninvasive pharmacologic (acetaminophen, nonsteroidal anti-inflammatory drugs, opioids, skeletal muscle relaxants, benzodiazepines,	Low back pain, back-specific and overall function, HRQoL, work disability, global improvement,	A systematic review of RCTs and systematic reviews published from through April 2015, updated	ACP grading system, adopted from the classification developed by the GRADE workgroup	Magnitude of effect determined as follows: For pain: • Mean post-treatment between-group difference on a VAS of 0 to 100	The systematic review and guideline underwent peer review and public comments from internal

Intended Users, Target Population	Intervention and Practice Considered	Major Outcomes Considered	Evidence Collection, Selection, and Synthesis	Evidence Quality Assessment	Recommendations Development and Evaluation	Guideline Validation
acute (< 4 weeks), subacute (4 to 12 weeks), and chronic (> 12 weeks) nonradicular low back pain, radicular low back pain, or symptomatic spinal stenosis	antidepressants, antiseizure medications, and systemic corticosteroids) and nonpharmacologic (psychological therapies, multidisciplinary rehabilitation, spinal manipulation, acupuncture, massage, exercise and related therapies, and various physical modalities) treatments for low back pain	back pain episodes, time between episodes, patient satisfaction, and AEs.	searches performed through November 2016		or equivalent: 5 – 10 points (small), > 10 – ≤ 20 points (moderate) • Mean between-group difference on a numerical rating scale of 0 to 10: 0.5 – 1.0 point (small), > 1.0 – ≤ 2.0 points (moderate) • Standardized mean difference of 0.2 – 0.5 (small effect), > 0.5 – ≤ 0.8 (moderate) For function: • Mean between-group difference on the ODI: 5 – 10 points (small), > 10 – ≤ 20 points (moderate) • Mean between-group difference on the RDQ: 1 – 2 points (small), > 2 – ≤ 5 points (moderate) • Standardized mean difference of 0.2 – 0.5 (small effect), > 0.5 – ≤ 0.8 (moderate) No large effects were found with any intervention.	and external stakeholders

ACP = American College of Physicians; AHRQ = Agency for Healthcare Research and Quality; GRADE = Grading of Recommendations Assessment, Development and Evaluation; HRQoL = Health related quality of life; ODI = Oswestry Disability Index; RDQ = Roland Morris Disability Questionnaire; VAS = visual analogue scale

Appendix 3: Critical Appraisal of Included Publications

Table 5: Strengths and Limitations of Systematic Reviews and Meta-Analyses AMSTAR[10]

Strengths	Limitations
Whale et al. 2019[16]	
• Overall, the review was conducted following Cochrane guidance to ensure the methodology was robust and systematic. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with predefined search strategy and no language restrictions applied and in multiple databases. In addition, a manual search from reference list of retrieved papers and review articles was also performed. The reporting of the search strategy followed the requirements of PRISMA statement. • Study selection was done by one reviewer, which was validated by another investigator, disagreements were resolved by discussion or consensus involving a third investigator. • Data extraction and validation were done independently by separate researchers. Investigators of the primary studies were contacted if disaggregated data were not reported; and studies were excluded if contact was unsuccessful. • The Cochrane risk of tool for RCTs was used to assess the risk of bias and quality of evidence, with information extracted and validated independently by separate researchers. However, no information was provided on how disagreements were resolved. • Study and patient characteristics of each included trial were provided with adequate details. • Opportunities for pooling data for meta-analysis were explored but limited because of heterogeneity in the content, duration, and intensity of the interventions, and conclusions were therefore based on narrative synthesis. Between-study heterogeneity was not assessed quantitatively as meta-analysis could not be conducted. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	• All included trials had high or unclear risk of bias for at least two domains in the Cochrane risk of bias tool. Overall, the included studies performed poorly in the domains of performance, detection, attrition, and reporting bias. • An assessment of publication bias was not done, and not justified.
Mariano et al. 2018	
• Overall, study methodology appeared to be appropriate and robust. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design.	• There was no information on whether data extraction and validation were done independently by separate researchers or not. • An assessment of methodological quality of the included studies was not made, although the authors stated an absence of widely accepted rating systems for prospective

Strengths	Limitations
• A systematic and comprehensive literature search was conducted, with predefined search strategy and in multiple databases. In addition, a manual search from reference list of retrieved papers and review articles was also performed. • Study selection was done independently by two reviewers, disagreements were resolved by discussion or consensus involving a third investigator. • Study and patient characteristics of each included trial were provided with adequate details. • Opportunities for pooling data for meta-analysis were explored but limited due to too few studies included, as well as heterogeneity in almost every key aspect—most notably in how the definition of condition, nature and frequency of the CBT intervention, length of follow-up, and the outcome measures. Conclusions were therefore based on narrative synthesis. Between-study heterogeneity was not assessed quantitatively as meta-analysis could not be conducted. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	studies as the reason for this. However, a qualitative analysis and summary was made for each study. • An assessment of publication bias was not done, and not justified.
Nicholls et al. 2018[5]	
• Overall, study methodology appeared to be appropriate and robust. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with predefined search strategy and in multiple databases. In addition, additional articles were identified through a hand search of relevant bibliographies. The reporting of the search strategy followed the requirements of PRISMA statement. • Study selection was done independently by two reviewers, disagreements were resolved by consensus. • The Cochrane risk of tool for RCTs was used to assess the risk of bias and quality of evidence, with information extracted and validated independently by separate researchers. Any discrepancies in bias-risk ratings were resolved by an independent third researcher. Most trials had generally low risk of bias across the six domains, with some reporting performance and detection bias to be high. • Three authors independently abstracted data from each article, and any discrepancies in data capture were resolved via consensus. • Study and patient characteristics of each included trial were provided with adequate details. • Pooling data for meta-analysis was not conducted due to heterogeneity in interventions, statistical analyses, follow-up periods, and outcomes assessed. Conclusions were therefore based on narrative synthesis. Between-study	• An assessment of publication bias was not done, and not justified.

Strengths	Limitations
heterogeneity was not assessed quantitatively as meta-analysis could not be conducted. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	
Silva Guerrero et al. 2018[14]	
• Overall, study methodology appeared to be appropriate and robust. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with predefined search strategy and no language restrictions applied and in multiple databases. In addition, a manual search from reference list of retrieved papers and review articles was also performed. The reporting of the search strategy followed the requirements of PRISMA statement. • Study selection was done in duplicate, disagreements were resolved by discussion or consensus involving a third reviewer. • Data extraction and validation were done independently by separate researchers. Disagreements were resolved by consensus and otherwise by consultations with a third reviewer. • Two reviewers independently assessed the risk of bias and quality of evidence using the Cochrane risk of bias tool for RCTs, with disagreements were resolved by discussion and consensus. All of the reviewed studies showed a high risk of performance bias, whereas up to a quarter of studies showed a high risk of attrition bias and other biases. • Study and patient characteristics of each included trial were provided with adequate details. • Statistical methods for data analysis appeared to be appropriate, between-study heterogeneity was assessed, subgroup analyses were done that were aligned with the objective of the systematic review. • Funnel plots were used to identify publication bias and assess reporting bias. Asymmetry of the funnel plots was assessed visually and using the Egger test when a minimum of 10 studies were included in the meta-analysis because the test power of fewer than 10 studies is usually too low to distinguish chance from real asymmetry. • The overall quality of the evidence for the main analyses was assessed using the GRADE approach. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	

Strengths	Limitations
Monticone et al. 2015	
• Overall, study methodology appeared to be appropriate and robust. • The scope of the systematic review was clear, with predefined inclusion/exclusion criteria for patient population, interventions, comparators, outcomes and study design. • A systematic and comprehensive literature search was conducted, with predefined search strategy and no language restrictions applied and in multiple databases. In addition, a manual search from reference list of retrieved papers and review articles was also performed. The reporting of the search strategy followed the requirements of PRISMA statement. • Study selection was done by 5 teams in duplicate, disagreements were resolved by discussion or consensus involving another reviewer. • Data extraction were done independently by separate researchers. Disagreements were resolved by consensus and in consultations with a third reviewer if disagreement persisted. • Two reviewers independently assessed the risk of bias and quality of evidence for each included RCT using the 12 criteria recommended by the Cochrane Back Review Group, which is an expansion of the criteria described in the Cochrane Handbook for Systematic Reviews of Interventions. Most of the included studies reported low or unclear risk of bias in almost all domains except for the domains associated with blinding. • Study and patient characteristics of each included trial were provided with adequate details. • Statistical methods for data analysis appeared to be appropriate, between-study heterogeneity was assessed, Subgroup analyses were planned but not conducted due to insufficient numbers of studies in each pairwise comparison. • The overall quality of the evidence for the main analyses was assessed using the GRADE approach. • Funnel plots were used to assess reporting bias when at least 10 studies were included in a meta-analysis and studies were not of similar size. Asymmetry of the funnel plots was assessed visually and using the Egger test Harbord method. • All authors and reviewers of the report declared no conflicts of interest. The source of funding for the report was provided.	

CBT = Cognitive-behavioral techniques; GRADE = Grading of Recommendations, Assessment, Development and Evaluation; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT = randomized controlled trial

Table 6: Strengths and Limitations of Clinical Studies using Downs and Black Checklist[11]

Strengths	Limitations
Mas et al. 2019[17]	
Reporting • The objective of the study, main outcomes, inclusion and exclusion criteria, interventions being compared, potential confounders, and main findings were clearly described. • The estimates of random variability (standard deviation/error or 95% confidence intervals) and exact P values were reported for the main outcomes. **External validity** • Participants in the trial were generally representative of the population from which they were recruited. Patients were excluded if they had more than 12 weeks of pain, history of prior LBP within 6 months, LBP that coexisted with cognitive impairment or psychiatric disorders, other causes of disability which impeded responding to the questionnaires, pregnancy and breastfeeding, physical problems in the preceding 3 months; and a diagnosis of fibromyalgia. These exclusion criteria do not appear to seriously affect the generalizability. **Internal validity** • Randomization method was reported, a reasonable justification was provided why a cluster randomization method was chosen i.e. to minimize contamination. • This was a single-blinded trial, with unblinded physician and blinded patients. Additionally, the analysts were blinded to the treatment allocation, adding to the strength of the findings. • The interventions were administered in a standardized manner, by allowing only one qualified psychologist and one physiotherapist with expertise in group interventions to implement the interventions in all healthcare centres. Further, data collection was done in a robust manner. • The outcomes measured were valid and reliable, or were widely accepted in the area of pain research. • The statistical tests were appropriate; in accordance with CONSORT guidelines. The effect of cluster was taken into account using appropriate analytical method. Multiple imputation method was used to address potential biases due to incomplete follow-up. The final models were adjusted for significant confounders and significant interaction variables. • There were no major imbalances between treatment arms in terms of baseline demographic and clinical characteristics, indicating randomization was preserved across clusters.	**Internal validity** • Control for multiple comparisons was not in place with a priori statistical hierarchy for secondary outcomes. • Some confounding factors were not accounted for in the model and may limit the conclusions of the study. For example, the differences in the profile of patients, since they were allocated by primary healthcare centres and socioeconomic status was not considered. • There were more healthcare centres allocated to the intervention group than the control group; however, the impact of this observation on the results is unclear. **Sample size/power** • Even though the authors conducted a power calculation, it is unclear if the trial had sufficient power to detect a clinically important effect for the primary endpoint. The authors estimated that 348 participants per intervention arm would allow detecting a mean difference of 2.5 points (SD 5.7) between the groups after accounting for 20% dropout. However, the number of patients recruited in the intervention and control group (n = 262 and 239, respectively) was well short of the target sample size. the number was even smaller at 3 and 12 months of follow-up.

CONSORT = Consolidated Standards of Reporting Trials; LBP = low back pain

Table 7: Strengths and Limitations of Guidelines using AGREE II[18]

Item	Guideline
	ACP guideline 2017[9]
Domain 1: Scope and Purpose	
1. The overall objective(s) of the guideline is (are) specifically described.	Yes
2. The health question(s) covered by the guideline is (are) specifically described.	Yes
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.	Yes
Domain 2: Stakeholder Involvement	
4. The guideline development group includes individuals from all relevant professional groups.	Unclear
5. The views and preferences of the target population (patients, public, etc.) have been sought.	Yes
6. The target users of the guideline are clearly defined.	Yes
Domain 3: Rigour of Development	
7. Systematic methods were used to search for evidence.	Yes
8. The criteria for selecting the evidence are clearly described.	Yes
9. The strengths and limitations of the body of evidence are clearly described.	Yes
10. The methods for formulating the recommendations are clearly described.	Yes
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	Yes
12. There is an explicit link between the recommendations and the supporting evidence.	Yes
13. The guideline has been externally reviewed by experts prior to its publication.	Yes
14. A procedure for updating the guideline is provided.	Yes
Domain 4: Clarity of Presentation	
15. The recommendations are specific and unambiguous.	Yes
16. The different options for management of the condition or health issue are clearly presented.	Yes
17. Key recommendations are easily identifiable.	Yes
Domain 5: Applicability	
18. The guideline describes facilitators and barriers to its application.	Not reported
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	Not reported
20. The potential resource implications of applying the recommendations have been considered.	Not reported

Item	Guideline
	ACP guideline 2017[9]
21. The guideline presents monitoring and/or auditing criteria.	Unclear
Domain 6: Editorial Independence	
22. The views of the funding body have not influenced the content of the guideline.	Yes
23. Competing interests of guideline development group members have been recorded and addressed.	Yes

Appendix 4: Main Study Findings and Authors' Conclusions

Table 8: Summary of Findings Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusion
Whale et al. 2019[16]	
Findings of each relevant trial up to month 3 are described below. **Music:** • 5 RCTs with 256 participants evaluated the effectiveness of music therapy for reducing acute postoperative pain following TKR. • One 2-arm study reported lower mean VAS pain scores at 3 hours (1.5 (SD 1.4) vs 3.9 (SD 3.4); p=0.01) and 24 hours (2.4 (SD 1.7) vs 4.1 (SD 2.9); p=0.04) post-surgery with patient-selected music played on headphones during surgery compared to white noise. • The remaining studies found no benefits of 20-minute 'easy listening' music, 30-minute soothing piano and Chinese violin music , 12 to 15-minute instrument only music with varying degrees of harmonicity and rhythmicity, and 5-minute music therapist-delivered live music compared to similarly timed quiet rest period, usual care, wearing headphones with no input, and no music, respectively; with respect to VAS or NRS pain score at various timepoints (minutes to days). **Guided imagery and music:** • One 2-arm RCT evaluated the effectiveness of guided imagery on outcomes post-surgery with 82 participants. • The study compared 19–21 minutes of audio-recorded guided imagery combined with soothing instrumental background music daily for 2 weeks before surgery and 3 weeks after surgery to a control group who received 17–21 min of commercially available spoken word audio recordings (eg, poetry, short stories, essays) at the same timepoints. However, no comparisons were made between trial arms even though pain was measured using the WOMAC pain score and VAS pain score up to 6 months post-surgery. **Hypnosis:** • One 3-arm RCT with 24 patients evaluated the effectiveness of a prerecorded hypnotic audio recording on pain outcome post-surgery. • There was a small difference in mean NRS ratings between patients who received 35-minute prerecorded hypnosis audio (listened to at least once pre-surgery and at least once 24 hours post-surgery) compared to those who received minimal treatment effect (psychoeducation, diaphragmatic breathing, relaxing music), and treatment as usual, no statistical comparison reported – mean NRS score at 72 hours: 1.77 vs 2.23 vs 2.59.	"*Due to the high heterogeneity of interventions and poor reporting of harms data, it was not possible to make any definitive statements about the overall effectiveness or safety of psychology interventions for pain outcomes after TKR*" pg 1

Main Study Findings	Authors' Conclusion
Progressive muscle relaxation with biofeedback: • One 2-arm RCT with 66 participants compared a multimodal combination of 3—minute training on biofeedback-assisted progressive muscle relaxation skills pre-surgery and 30-minute sessions of continuous passive motion therapy twice a day for 5 days post-surgery to standard continuous passive motion therapy. • The former intervention showed a statistically significant lower NRS pain scores compared with the control group P < 0.001). **Pain coping skills program:** • One 3-arm RCT with 402 patients evaluated the effectiveness of a pain coping skills program to arthritis education and to usual care. The program comprised of eight 50-minute sessions over a 2-month period beginning 2 weeks pre-surgery and ending 6 weeks after surgery and included sessions on cognitive restructuring, thought identification and challenging, self-calming and relaxation techniques, and activity management; delivered in-person and via telephone. • At baseline and up to 12 months post-surgery, no differences were found in mean WOMAC pain treatment scores or group-by-time interaction. **Enhanced postoperative recovery using motivational interviewing:** • One 2-arm RCT with 308 participants evaluated the effectiveness of an enhanced postoperative recovery program (comprising 10 telephone calls with a navigator over a 6-month postoperative period aimed at identifying postsurgical objectives and improving self-efficacy using motivational interviewing) compared to usual care (including inpatient physiotherapy and outpatient physiotherapy after discharge) to improve postoperative functional status. • At baseline and up to 6 months post-surgery, no differences were found in mean WOMAC pain treatment scores between interventions. **CBT programs:** • 2 RCTs with 150 participants evaluated the effectiveness of CBT-based programs. • One 2-arm RCT with 50 participants compared CBT-based program (up to ten 1-hour sessions for reducing anxiety and depression) to standard care. At 4 months post-surgery, no differences were found in mean WOMAC pain treatment scores between the interventions.	
Mariano et al. 2018[2]	
• 1 RCT with 93 patients with subacute LBP compared CBP (individual education, coping strategies, reassurance (minimum 2 sessions) with exercise and control. Results	"*Five of the six showed significant improvements associated with CBT, but the heterogeneity of the studies prevented quantitative comparisons*" pg 1

Main Study Findings	Authors' Conclusion
showed improved disability in CBT group vs. control (P = 0.02). However, no indication which outcome favored CBT (VAS, RMDQ), SF-36, Cantril's Ladder Scale. • 1 RCT with 64 patients with subacute LBP compared the clinical benefits of CBP (9 individual biopsychosocial treatment 3×/week and 12 relaxation sessions, 4×/week) and control. Results showed reduced pain intensity (VAS) for CBT group vs. control at 6 months (all P ≤ 0.005). • 1 RCT with 67 patients with subacute LBP compared the clinical benefits of CBP (6 sessions of Individual, custom program including problem solving, coping skills, relaxation) and control. Results showed CBT group had greater "recovery" (resolution of pain [DDS] and restoration of function [SIP] at 6 months) if completed 4 (P = 0.02) or 6 (P = 0.002) sessions. However, no results provided for health status (QWB). • 1 RCT with 38 patients with subacute LBP compared the clinical benefits of CBP (Individual psychotherapy, minimum 7 sessions, 1×/week) and control. However, none of the reported outcomes were relevant for this report. • 1 RCT with 701 patients with a mix of subacute and chronic LBP compared the clinical benefits of CBP (6 sessions of group therapy addressing "behaviors and beliefs about physical activity and avoidance of activity") and control. Results showed CBT group had significantly improved RMDQ, von Korff score, SF-12 physical component subscale at 3, 6, and 12 months (all P < 0.032); unknown how much of this effect was driven by chronic LBP participants. • 1 RCT with 216 patients with a mix of subacute and chronic LBP compared the clinical benefits of CBP (4 sessions of "Cognitive educational program") and control. Results showed CBT group was not significantly different than control with respect to RMDQ, VAS, and EQ-5D.	"*CBT has not been adequately studied as a potential early intervention treatment for sALBP patients. None of the six identified papers studied US civilians or leveraged innovations such as teletherapy—able to reach patients in remote or underserved areas—underscoring critical gaps in current back pain treatment. Given the severity of the US opioid epidemic, nonpharmacologic options such as CBT should be rigorously explored in the sALBP population*" pg 1
Nicholls et al. 2018[5]	
Findings of each relevant trial up to month 3 are described below. Trial 1: • In a Swedish trial consisting of 107 patients undergoing lumbar fusion surgery, the clinical effectiveness of three 90-minute in-person PMT sessions (SOC+ CBT-informed physiotherapy intervention including behavioral self-management, cognitive restructuring, relaxation training, and motivational interviewing) and usual care (physiotherapist-instructed inpatient exercise therapy program and 20-minute exercise instruction at discharge) were compared. • At 3 months, Disability (ODI): –9.7 (–15.8 to –3.6, P=0.002), d=1.2. Back pain (VAS): –11.7 (–19 to –4.3, P=0.002), d=1.45. Trial 2:	"*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*" pg 1

Main Study Findings	Authors' Conclusion
• In an American trial consisting of 86 patients undergoing laminectomy, the clinical effectiveness of CBPT (6 weekly 30-minute telephone-based CBPT sessions consisted of behavioral self-management, problem-solving, cognitive restructuring, and relaxation training + education program) and education (education program consisting of physical therapy, stress management, sleep hygiene, energy management, communication with health providers, injury prevention) were compared at 3 months. • BPI back pain: –0.88 (–1.5 to –0.25, P=0.007), d=0.62 • BPI leg pain: –1.2 (–2.1 to 0.34, P=0.007), d=0.62 • BPI interference: –1.5 (–2.4 to –0.57, P=0.002), d=0.72 • ODI: –9.8 (–15.3 to –4.4, P<0.001), d=0.79 • Five-chair-stand test: –7 (–13.7 to –0.37, P=0.04), d=0.49 • TUG test: –1.6 (–3.3 to 0.19, P=0.08); d=0.41 • 10 m walk test: 0.1 (–0.14 to 0.21, P=0.08), d=0.41–0.49 Between-group differences in score reduction • BPI back pain: –0.85 (–1.4 to –0.25, P=0.006, R^2=0.64) • BPI leg pain: –1.1 (–1.9 to –0.2, P=0.009, R^2=0.447) • BPI pain interference: –1.3 (–2.1 to –1.4, P=0.005, R^2=0.49) • ODI: –9.4 (–14.9 to –4, P=0.001, R^2=0.59) • Five-chair-stand test: –4.3 seconds (–7.7 to –0.82, P=0.02, R^2=0.52) • TUG test: –1.8 seconds (–3.2 to –0.16, P=0.02, R^2=0.62) • 10 m walk test: m/s (0.008 to 0.18, P=0.07, R^2=0.33) Trial 3: • In an Italian trial consisting of patients undergoing lumbar spinal fusion, the clinical effectiveness of CBT + exercise (two 1-hour psychologist-delivered individual CBT sessions per week over 4 weeks targeting catastrophizing and kinesophobia, in addition to exercise program) and exercise (five 90-minute physiotherapist-led exercise sessions per week for 4 weeks) were compared at 4 weeks. • At 4 weeks post-surgery, all group effects and time effects significant at P<0.001 level. Group × time effects Effect size corresponds to difference between groups • ODI: F=20.37, P<0.001, d=0.8 • NRS back: F=40.87, P<0.001, d=1.13 • NRS leg: F=12.32, P<0.001, d=0.62 • SF-36-BP: F=12.25, P<0.001, d=0.62 Trial 4: • In an Danish trial consisting of patients undergoing lumbar spinal fusion patients, the clinical effectiveness of CBT (four 3-hour group CBT sessions in addition to usual care) and usual care (information on operation, anesthesia, medication, PO rehab, physical restrictions after surgery) were compared at 7 days post-surgery.	

Main Study Findings	Authors' Conclusion
• PO mobility improved on day 3: Walk, P=0.02, Rise and sit from chair, P=0.0017, Get in and out of bed, P=0.0017). However, pain and rescue analgesic use were not statistically significant. Trial 5: • In the same Danish trial described above, a longer 3-month follow-up showed an improvement in disability (ODI), P=003; however, back pain and leg pain were not.	
Silva Guerrero et al. 2018[14]	
Low to high quality evidence demonstrated small to medium effects for some psychological outcomes at short-term and long-term follow-ups. Only short-term pain and disability-related outcomes are reported below. Pain outcomes: • There was moderate quality evidence that physiotherapist delivered psychological intervention combined with physiotherapy decreased pain in the short-term (26 studies, n = 3024, MD=− 0.37; 95% CI: −0.65 to −0.09; I^2= 63%; P=0.008), although the effect size was small. • No statistically significant difference was found for pain in the subgroup of patients with LBP only at the short-term (15 studies, n=1921; MD=−0.33; 95% CI: −069 to 0.03; I^2=67%; P=0.07). • No statistically significant difference was found for pain in the subgroup of patients with WAD and neck pain only at the short-term (7 studies, n=683; MD=−0.27; 95% CI, −0.81 to 0.26; I^2=55%; P=0.31). • No statistical difference was found for self-efficacy in the subgroup of patients with OA-related pain at short-term follow-up (4 studies, n=420; MD=−0.73; 95% CI: −1.61 to 0.15; I^2= 70%; P= 0.10) Disability Outcomes: • Low-quality evidence showed that psychological interventions combined with physiotherapy was associated with a statistically significant but small improvement in disability compared with physiotherapy alone or usual care in the short term (29 studies, n=4249; SMD=−0.14; 95% CI: −0.26 to −0.01; I^2=72%; P=0.03). • No statistically significant difference was found for pain in the subgroup of patients with LBP only at the short-term (17 studies, n= 2298; SMD=−0.12; 95% CI, −0.31 to 0.07; I^2= 80%; P= 0.21). • No statistically significant difference was found for pain in the subgroup of patients with WAD and neck pain only at the short-term (9 studies, n= 1605; SMD=−0.14; 95% CI, −0.32 to 0.04; I^2=60%; P= 0.12). • No statistical difference was found for self-efficacy in the subgroup of patients with OA-related pain at short-term follow-up (3 studies, n=346; SMD=−0.20; 95% CI, −0.41 to 0.01; I^2=0%; P= 0.06).	"*The results indicate that psychological interventions delivered by physiotherapist show promise to improve health outcomes, particularly psychological outcomes, in musculoskeletal pain conditions*" pg 1 "*The results of the review of these 34 studies indicated that combined physiotherapy and psychological interventions delivered by physiotherapists shows promise to improve health outcomes, particularly psychological outcomes, in musculoskeletal pain conditions*" pg18

Main Study Findings	Authors' Conclusion
Monticone et al. 2015[15]	
• Based on meta-analysis of 2 studies comparing the effects of CBT versus other interventions (information booklet, didactic discussion, manual therapy) at short-term (6–13 weeks) and long-term follow-up (52 weeks), there was low quality evidence that CBT was statistically significantly better than other interventions for improving subacute neck pain (SMD -0.24, 95% CI: -0.48 to 0.00, I^2 = 7%, P = 0.05), while no difference was found in terms of disability (SMD -0.12, 95% CI: -0.36 to 0.12, I^2 = 0%, P = 0.31) • One of the two included RCTs showed that the CBT group outperformed the no treatment group in terms of pain and disability (mean pain severity: possible range 0 to 6, MD -0.80, 95% CI: -1.27 to -0.33); Neck Disability Index MD -5.80, 95% CI: -10.52 to -1.08). • The remaining RCT reported that CBT was better than manual therapy at improving pain and disability (Numerical Rating Scale MD 0.99, 95% CI 0.15 to 1.83; Neck Disability Index MD 2.42, 95% CI 0.52 to 4.32). For other outcomes (such as Tampa Scale of Kinesiophobia or the Pain Coping and Cognition List), there was no significant difference between groups.	"*For patients with subacute NP, CBT was significantly better than other types of interventions at reducing pain at short-term follow-up, while no difference was found for disability and kinesiophobia*" pg 5

ACT = Acceptance and commitment therapy; BPI = Brief Pain Inventory; CAS = Cumulated Ambulation Score; CBT = Cognitive-behavioral techniques; CLS = Cantril's Ladder Scale; DDS = Descriptor Differential Scale; EQ-5D = Euro Quality of life-5 Dimension; HRQoL = Health related quality of life; LBP = low back pain; MPQ = McGill Pain Questionnaire; NDI = Neck Disability Index; NRS = Numeric Rating Scale; OA = Osteoarthritis; ODI = Oswestry Disability Index; RCT = randomized controlled trial; TKR = total knee replacement; QWB = Quality of Well-Being; RMDQ = Roland Morris Disability Questionnaire; SAE = serious adverse events; SIP = Sickness Impact Profile; SF-12/36 = Short Form 12/36; VAS = visual analogue scale; WAD = Whiplash-associated disorders; WOMAC = Western Ontario and McMaster Universities Osteoarthritis Index

Table 9: Summary of Findings of Included Primary Clinical Studies

Main Study Findings	Authors' Conclusion
Mas et al. 2019[17]	
Disability: • The biopsychosocial multidisciplinary intervention group showed a statistically improvement in the adjusted analysis of the RMDQ outcome compared to the control group at 3 months (− 1.33 points, 95% CI: − 2.22 to − 0.45, p = 0.005) and at 12 months (− 1.11 points, 95% CI: − 2.08 to − 0.13, P = 0.027). • An MCID (2.5 points from baseline as reported by the authors) was achieved in both groups, with a difference over 3.5 points in the intervention group compared with baseline at each time-point (3.8 RMDQ points at 3 months and 5.1 RMDQ points at 12 months). **Pain:** • With respect to the total intensity score of MPQ, a marginal difference was observed in the intervention group compared to the control group (− 0.69 points; 95% CI: − 1.41 to 0.02; p = 0.058) at 12 months, but not at 3 months (−0.49 points; 95% CI: −1.39 to 0.42; p = 0.294). • The intervention group presented a statistically significant differences compared to the control group at 3 months for current intensity score (− 0.32 points; 95% CI: − 0.63 to − 0.02; p = 0.040) and for VAS score (− 0.77 points; 95% CI: − 1.53 to − 0.01; P = 0.046). However, no statistically significant differences were seen for either endpoints at 12 months (−0.18 points; 95% CI: − 0.43 to 0.08; p = 0.162 for current intensity score and (−0.27 points; 95%CI: − 0.88 to 0.34; P = 0.374 for VAS). **Quality of life:** • The outcome of SF-12 increased in both groups during the follow-up period, but no statistically significant differences between groups were observed on the physical domain (0.55 points, 95% CI: − 1.19 to 2.29, P = 0.520 and 0.53 points, 95% CI: −1.20 to 2.27, P = 0.532 at 3 and 12 months, respectively) and mental health domain (2.56 points, 95%CI: − 0.33 to 5.45, P = 0.082 and 1.48 points, 95% CI: −0.86 to 3.83, P = 0.206 at 3 and 12 months, respectively).	"*A multidisciplinary biopsychosocial intervention in a working population with non-specific subacute LBP has a small positive impact on disability, and on the level of pain, mainly at short-term, but no difference on quality of life*" pg 2 "*The main conclusion of this study is that a multidisciplinary biopsychosocial intervention in a working population with non-specific subacute LBP has a small positive effect on disability and intensity of pain. Although greater in the intervention group, minimal clinically important differences were achieved in both groups. The results did not show any differences on quality of life*" pg 8

CBT = cognitive-behavioural therapy; MCID = minimal clinically important difference; MPQ = McGill Pain Questionnaire; LBP = low back pain; RCT = randomized controlled trial; RMDQ = Roland Morris Disability Questionnaire; SF-12 = Short Form 12-Item; VAS = visual analogue scale

Table 10: Summary of Recommendations in Included Guidelines

<table>
<tr><th>Recommendations</th><th>Strength of Evidence and Recommendations</th></tr>
<tr><td colspan="2">American College of Physicians, 2017[9]</td></tr>
<tr><td>No recommendation based on psychotherapies given for acute/subacute pain.

“Given that most patients with acute or subacute low back pain improve over time regardless of treatment, clinicians and patients should select nonpharmacologic treatment with superficial heat (moderate-quality evidence), massage, acupuncture, or spinal manipulation (low-quality evidence). If pharmacologic treatment is desired, clinicians and patients should select nonsteroidal anti-inflammatory drugs or skeletal muscle relaxants (moderate-quality evidence).” pg 1

Grade: strong recommendation</td><td>Criteria for strength of definition:

<table><tr><td>Quality of Evidence</td><td colspan="2">Strength of Recommendation</td></tr><tr><td></td><td>Benefits Clearly Outweigh Risks and Burden or Risks and Burden Clearly Outweigh Benefits</td><td>Benefits Finely Balanced With Risks and Burden</td></tr><tr><td>High</td><td>Strong</td><td>Weak</td></tr><tr><td>Moderate</td><td>Strong</td><td>Weak</td></tr><tr><td>Low</td><td>Strong</td><td>Weak</td></tr></table></td></tr>
</table>

ACP = American College of Physicians; AHRQ = Agency for Healthcare Research and Quality; GRADE = Grading of Recommendations Assessment, Development and Evaluation; HRQoL = Health related quality of life; ODI = Oswestry Disability Index; RDQ = Roland Morris Disability Questionnaire; VAS = visual analogue scale

Appendix 5: Overlap between Included Systematic Reviews

Table 11: Overlaps between Included Systematic Reviews

Primary Study Citation	Systematic Review Citation				
	Whale et al. 2019[16]	Mariano et al. 2018[2]	Nicholls et al. 2018[5]	Silva Guerrero et al. 2018[14]	Monticone et al. 2015[15]
Allred et al 2010	X				
Cai et al 2018	X				
Chen et al 2015	X				
das Nair et al 2018	X				
Jacobson et al 2016	X				
Finlay et al 2016	X				
Lee et al 2019	X				
Leonard 2019	X				
Losina et al 2016	X				
Riddle et al 2019	X				
Simock et al 2008	X				
Wang et al 2015	X				
Abbott et al 2010			X		
Archer et al 2016			X		
Doering et al 2016			X		
Monticone et al 2014			X		
Rolving et al 2016			X		
Rolving et al 2015			X		
Storheim et al 2003		X			
Schiltenwolf et al 2006		X			
Lindell et al 2008		X			
Slater et al 2009		X			
Lamb et al 2010		X			
Werner et al 2016		X			
Asenlof et al 2005				X	
Basler et al 2007				X	
Bennell et al 2016				X	

Primary Study Citation	Systematic Review Citation				
	Whale et al. 2019[16]	Mariano et al. 2018[2]	Nicholls et al. 2018[5]	Silva Guerrero et al. 2018[14]	Monticone et al. 2015[15]
Bring et al 2015				X	
Critchley et al 2007				X	
Friedrich et al 1998				X	
George et al 2003				X	
George et al 2008				X	
Gustavsson et al 2010				X	
Gustavsson et al 2011				X	
Hay et al 2005				X	
Hill et al 2011				X	
Hunt et al 2013				X	
Johnson et al 2007				X	
Johstone et al 2002				X	
Lamb et al 2012				X	
Lee et al 2013				X	
Ludvigsson et al 2015 and 2016				X	
Macedo et al 2012				X	
Magalhaes et al 2013				X	
Moffett et al 2005				X	
Moffett et al 2006				X	
Monticone et al 2012				X	
Murray et al 2015				X	
Patel et al 2014				X	
Pool et al 2010				X	X
Saw et al 2016				X	
Soderlund et al 2001 and 2007				X	
Staal et al 2004				X	
Steenstra et al 2006				X	

Primary Study Citation	Systematic Review Citation				
	Whale et al. 2019[16]	Mariano et al. 2018[2]	Nicholls et al. 2018[5]	Silva Guerrero et al. 2018[14]	Monticone et al. 2015[15]
Van der Roer et al 2008				X	
Veenhof et al 2006				X	
Vibe Fersum et al 2013				X	
Vong et al 2011				X	
Vonk et al 2009				X	
Robinson et al 2013					X

Appendix 6: Additional References of Potential Interest

Studies with mixed population of acute and chronic musculoskeletal pain where results were not presented separately

Randomized Controlled Trial of Brief Mindfulness Training and Hypnotic Suggestion for Acute Pain Relief in the Hospital Setting.
Garland EL, Baker AK, Larsen P, Riquino MR, Priddy SE, Thomas E, Hanley AW, Galbraith P, Wanner N, Nakamura Y. J Gen Intern Med. 2017 Oct;32(10):1106-1113. doi:10.1007/s11606-017-4116-9.[19]

A Brief Mindfulness Intervention for Medically Hospitalized Patients with Acute Pain: A Pilot Randomized Clinical Trial.
Miller-Matero LR, Coleman JP, Smith-Mason CE, Moore DA, Marszalek D, Ahmedani BK. Pain Med. 2019 Nov 1;20(11):2149-2154. doi: 10.1093/pm/pnz082.[6]

Systematic review used as the evidence base for the ACP guideline

Nonpharmacologic Therapies for Low Back Pain: A Systematic Review for an American College of Physicians Clinical Practice Guideline.
Chou R, Deyo R, Friedly J, Skelly A, Hashimoto R, Weimer M, Fu R, Dana T, Kraegel P, Griffin J, Grusing S, Brodt ED. Ann Intern Med. 2017 Apr 4;166(7):493-505. doi: 10.7326/M16-2459. Epub 2017 Feb 14.[8]