The effectiveness of peer support interventions for community-dwelling adults with chronic musculoskeletal pain: a systematic review and meta-analysis of randomised trials.

WILSON, M.V., BRAITHWAITE, F.A., ARNOLD, J.B., CROUCH, S.M., MOORE, E., HEIL, A., COOPER, K. and STANTON, T.R.

2024

INTRODUCTION

Chronic musculoskeletal (MSK) pain refers to pain experienced in the muscles, joints, bones, or surrounding structures that persists for longer than three months [69]. MSK conditions, estimated to affect 30% of the population [19], are the most common reason for chronic pain [94], and result in physical limitations, social isolation, and psychological distress [93].

Healthcare systems world-wide grapple with the escalating burden of chronic MSK pain. Health professionals face mounting challenges in addressing patient needs due to time constraints, limited resources and competing demands [26,29,62,63,68]. Many individuals with chronic MSK conditions do not have access to adequate subsidised healthcare and cannot self-fund (high costs) [33]. They may attempt to self-manage their condition, although inadequate access to support and guidance [6,83] often fosters reliance on low-value treatments (e.g., opioids) [30,46]. Evidence-based, accessible, and high-value adjunct healthcare options within current systemic constraints are needed.

People with lived experience of pain (i.e., peers) may be well-placed to offer support and guidance for others with chronic MSK pain. Peer support refers to "the giving of assistance and encouragement by an individual considered equal" [25 p323] as part of a created network or intervention. Here, 'equal' denotes individuals having the same health condition. Peer support encompasses emotional (listening, reflecting, empathy, reassurance), informational (advice, knowledge), and appraisal (affirmation, encouragement) support [26]. Peer support differs from informal support groups as 'peers' receive training to deliver the intervention [26].

Peer support holds potential to augment high-value components of chronic MSK pain care [27,67], such as self-management, which involves navigating the symptoms and adjustments associated with living with a chronic condition [10]. Patients may be more trusting and receptive to support and education from peers with analogous experiences, imparting greater benefit than general health guidance provided by health professionals [24]. Lack of 'shared identity' (via lived experience) can heighten scepticism when patients receive advice from someone who has not been in the same position as themselves [80]. This may reduce trust and rapport, creating a barrier to education and self-management [31,34,77]. While health professionals play a vital role in educating and supporting patients' active self-management [83], implementation of peer support may improve engagement, increase acceptability, and

alleviate healthcare system strain, particularly given that peer support can take varied forms (one-to-one, group, face-to-face, online, telephone, hybrid) [62].

A previous review (2014) found limited evidence (n=5 studies) for the effectiveness of peer support for adults with chronic non-cancer pain [23], and its wide scope (including all chronic non-cancer pain) resulted in high participant/outcome heterogeneity, preventing meta-analysis. While numerous clinical trials have been published since, warranting an update, focussing of the review question holds clinical relevance. Peer support interventions for non-MSK pain (e.g., AIDs) may have important differences to those for MSK pain (e.g., does pain reflect disease progression?), and differing prognostic trajectories and treatment response between MSK/non-MSK populations can muddy interpretation of clinical outcomes. Therefore, we focussed our review, aiming to summarise and critically appraise available evidence on the effectiveness of peer support interventions for people with chronic MSK pain.

METHOD

The specific research question of our review was: Are peer support interventions more effective than usual care, waitlist control and/or other active interventions, in reducing pain and health service utilisation, and enhancing function, quality of life, self-efficacy, self-management and perceived social support in community-dwelling adults with chronic MSK pain? This review is reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [73] and was prospectively registered on PROSPERO (CRD42022356850), with amendments tracked (see Appendix 1). All stages of the review were conducted by two independent reviewers. Any disagreements were resolved via discussion or, if needed, consultation with a third reviewer. Study authors were contacted a maximum of three times via email to: provide full texts; clarify eligibility; or provide missing/unamenable data. If data were irretrievable, studies were included in descriptive synthesis only. All data collection forms, extracted data, and analysis files are freely available on Open Science Framework.

Identification and selection of studies:

Study eligibility criteria are provided in Table 1. Peer support interventions were required to involve formal support provided and received by people with the same condition [26], in which peers delivering the intervention had received training. Studies in which peer mentors

were not trained ('lay-leaders') [26] or extensively trained ('paraprofessionals') [26] were excluded. Studies with peer support included as the predominant part of a multi-component intervention were included to enhance ecological validity of our findings given this is a common real-world scenario. Feasibility randomised controlled trials (RCTs) were eligible for inclusion. There were no limits on publication period, language, or publication status.

INSERT TABLE 1

A systematic search of MEDLINE, Embase, Emcare, PsycINFO (via OVID), CINAHL (EBSCO), The Cochrane Library (Wiley), and Scopus (Elsevier) was undertaken from inception to January 31, 2023. Keywords and database-specific subject headings (e.g., MeSH) relating to peer support, self-management, and chronic pain were used. The search strategy was developed in consultation with an academic librarian (see Appendix 2 for full search strategy). The terms 'lay-leader' and 'lay-led' were included given their historical use to describe peer support. A grey literature search was conducted, including: Clinical Trial Registries (ClinicalTrials.gov; Australian New Zealand Clinical Trails Registry); a thesis database (ProQuest Theses and Dissertations); and Google Scholar (two keyword strings were used; first 20 results from each exported). Reference lists of all included studies were hand searched for additional relevant studies, with citation tracking conducted on included studies using Web of Science (Clarivate, Philadelphia, USA) to identify any further relevant studies.

Search results were exported to EndNote (v.20.4, Clarivate, Philadelphia, USA) for removal of duplicates, and then exported to Covidence (Veritas Health Innovation Ltd., Melbourne, Australia), where further duplicates were removed. Following title and abstract screening, the full texts of potentially relevant studies were retrieved and then assessed for inclusion. Authors of trial protocols were contacted to determine data availability. Google Translate was used for non-English publications, with translations cross-checked for accuracy by fluent speakers of each language.

Data extraction:

A customised, piloted data extraction form was used to capture data on: publication details (authors, year); study design; recruitment; participant characteristics (age, gender, ethnicity, education, employment status, diagnosis, condition duration); intervention and control details (mode, format, duration, frequency, content); sample size and dropouts; study outcomes

(outcome measures, follow-up timepoints, measures of central tendency [mean, median, mean difference], and measures of dispersion [SD, IQR, standard error, CI]); and intervention adherence/attendance (when available). Data extraction related to social determinants of health was informed by the PROGRESS-Plus Framework by Cochrane [71].

Risk of bias (RoB) and evidence quality assessment:

The Cochrane RoB assessment tool for randomised trials (version 1) [39] was used to determine study-specific RoB across the following domains: randomisation; concealment; blinding; and detection; attrition; reporting; and 'other' biases. A lack of blinding of participants, personnel and of outcome assessment may have influenced outcomes. Consequently, blinding of participants and personnel was marked as 'high' risk for no, incomplete or broken blinding (including waitlist control comparators), and 'unclear' when there was insufficient information available for judgement. For blinding of outcome assessment (detection bias), all self-reported clinical outcomes were considered 'high' risk if participants were unblinded. The GRADE approach [8] was used to evaluate the certainty of available evidence for each comparison/outcome, based on the following domains: risk of bias; inconsistency; indirectness; imprecision; or publication bias (see Appendix 3).

Data analysis:

Studies were grouped by control comparison (usual care, waitlist control, or active control) and outcome (e.g., pain, self-efficacy), with sub-grouping based on follow-up timepoint: short-term (0-3 months); medium-term (4-9 months); and long-term (>9 months), consistent with prior literature [17]. Given wide variation in outcomes evaluated by included studies, a pragmatic decision was made to focus analyses on the primary outcomes delineated in our protocol, rather than reporting all available outcomes. These outcomes were: pain intensity, self-efficacy, function, health service utilisation, quality of life, social support, and self-management.

When two or more studies evaluated a similar outcome and used a similar control condition, meta-analyses were considered. Pooling was undertaken using RevMan Web (v5.4, Cochrane, London, UK) generic inverse-variance random effects models to calculate Mean Differences (MD) and 95% CIs. If studies used differing measures for the same outcome, then Standardised Mean Differences (SMDs) were calculated using Hedges' g (0.2 small, 0.5 moderate, and 0.8 large effect) [39]. To aid interpretation, SMDs were back-converted to the

original units of measurement of a commonly known outcome measure using the SD from the largest, most unbiased, and representative study among those pooled [39]. Post-intervention data were used for meta-analyses; if unavailable, within-group change from baseline was used [39]. Cochrane recommended methods were used to transform data into formats amenable to pooling (e.g., from change score to timepoint data), and when required, correlation coefficients were used/calculated from analogous studies [39]. If standard deviations (SDs) were unavailable, data were imputed using established methods [39]. Chi-squared (χ^2) and I-squared (I^2) statistics were used to assess heterogeneity, with $p<0.10$ interpreted as statistically significant and $I^2>50\%$ as substantial heterogeneity [39]. When meta-analyses included ≥10 studies, publication bias was assessed using visual inspection of funnel plots for asymmetry [85]. Where pooling was not possible, narrative synthesis was undertaken.

When multiple measures were used to assess the same outcome in a study, the outcome best aligned with that of other studies in the pooled analysis was chosen. When studies had multiple follow-up timepoints falling within our specified subgroups (e.g., short-term), timepoints most consistent with the other studies in each respective meta-analysis were used, with non-pooled timepoint results described narratively. The *function* outcome domain considered measures of either impaired physical function or disability. Studies reporting on health service utilisation were pooled based on country to account for differing health systems.

Five types of sensitivity analyses were conducted: (1) when correlation coefficients were used for data imputation, sensitivity analyses used correlation coefficient variation of ±0.1, consistent with best practice [39]; (2) when SDs were imputed, sensitivity analyses removed those studies, as well as imputed highest and lowest available SDs from studies in the pooled comparison; (3) when peer support was part of a complex multi-component intervention (i.e., unable to determine unique effect of peer support), sensitivity analyses removed these studies; (4) when studies included more than one peer support group, sensitivity analyses used the alternate peer group data (versus control); (5) when studies used multiple measures for the same outcome, sensitivity analyses used the alternate outcome measure data.

RESULTS

Flow of studies through review:

A total of 16,445 records were identified via database searches. Following removal of duplicates, 7684 unique records underwent title/abstract screening, with 578 full texts undergoing formal eligibility assessment. From this, 562 records were excluded, resulting in 16 records included via database searches (Figure 1). Grey literature searches identified 4990 records. Following removal of duplicates, 3936 records were screened, with 354 full texts assessed for eligibility, which resulted in 13 new records included via grey literature searches (Figure 1). Together, a total of 29 records describing 24 studies were included. Of sixty authors contacted, concerning 49 records, 41 provided additional information and/or data.

INSERT FIGURE 1

Characteristics of studies:

Study characteristics are summarised in Table 2. Studies were published between 1985 and 2021, recruiting a total of 6202 participants with chronic MSK pain (mean 69% women, range 18-91%; mean ages ranging 42.0 to 77.6 years). Studies were conducted across eight countries, with half undertaken in the United States (12/24). One study was a feasibility RCT [4].

Chronic MSK pain conditions studied included: osteoarthritis (OA; n=10 studies); arthritis (unspecified; n=6); rheumatoid arthritis or 'rheumatic disorders' (n=5); upper body or back pain (n=3), chronic MSK pain (unspecified; n=3), fibromyalgia (n=2), and ankylosing spondylitis (n=1). Four studies recruited mixed chronic MSK populations [15,47,57,58] and one study recruited an arthritis subsample as part of a larger study of many chronic conditions [59]. Thirteen studies (54%) reported participant ethnicity, the majority with predominately Caucasian participants. Education levels were reported in all but four studies, displaying significant variability.

INSERT TABLE 2

Intervention characteristics:

Intervention and control condition characteristics are summarised in Table 3. Most peer support interventions used in-person group sessions (19/24, 79%). Ten studies (10/24, 42%) evaluated the Arthritis Self-Management Program (ASMP), a community-based program promoting arthritis management [55]. Another three (3/24, 13%) evaluated the more generic

Chronic Disease Self-Management Program (CDSMP), which caters to various chronic conditions [55]. One study featured two peer-led intervention groups: one participant-only and another including participants alongside spouses [60]. Intervention duration ranged between two weeks [15] and 12 months [50], with the majority lasting six weeks (14/24, 58%). Adherence or attendance rates, reported by 71% studies, varied considerably (Table 3). Active control interventions included professionally led self-management courses, personalised mailed self-management programs, and professionally led tailored exercise groups. Two of the peer support interventions (2/24, 8%) actively involved individuals with chronic pain during the design stage of the intervention [4,87].

INSERT TABLE 3

RoB and GRADE outcomes:

Most RoB domains were rated high or unclear (Figure 2). While randomisation methods had predominantly low RoB (63%), allocation concealment was unclear in most studies (67%) and blinding of participants/personnel was unclear or high risk for all studies. Given self-reported outcomes, high risk of detection bias was present across all studies. Attrition bias was variable, with large variation in drop-out rates and missing data imputation methods. Few studies had pre-registered protocols, making reporting bias unclear in most cases (63%). GRADE evaluation showed that all outcomes had low to very low certainty evidence (see Appendix 3).

INSERT FIGURE 2

Outcomes:

Twenty-two studies (25 of 29 records) could be pooled; the remaining are provided via narrative synthesis [15,43,78,79]. Table 4 provides an overall summary of meta-analytic findings. Unpooled data can be found in Appendix 4.

INSERT TABLE 4

Effect of Peer Support Intervention on Pain Intensity (0-100 scale):

Versus usual care (N=9; Figure 3a)

Pooling showed no significant effect of peer support on pain in the short-term (MD: -1.58 [95%CI -4.83, 1.66], P=0.34, I^2=0%, N=5), although peer support interventions resulted in a significantly greater reduction in pain relative to usual care in the medium-term (MD: -3.48 [95%CI -6.61, -0.35], P=0.03, I^2=80%, N=6) and long-term (MD: -1.97 [95%CI -3.53, -0.42], P=0.01, I^2=0%, N=7). Heterogeneity was high in the medium-term.

Versus waitlist control (N=8; Figure 3b)

Pooling showed no effect of peer support on pain in the short-term (MD: 2.00 [95%CI -5.12, 9.12], P=0.58, I^2=0%, N=2) or the medium-term (MD: -2.90 [95%CI -6.62, 0.81], P=0.13, I^2=57%, N=7) when compared to waitlist controls. No studies evaluated long-term effects on pain.

Versus active control (N=7 studies; Figure 3c)

Pooling showed no effect of peer support for pain in the short-term (MD: 4.98 [95%CI -0.08, 10.04], P=0.05, I^2=15%, N=3), medium-term (MD: 1.90 [95%CI -1.79, 5.59], P=0.31, I^2=0%, N=3), or long-term (MD: 2.94 [95%CI -0.01, 5.90], P=0.05, I^2=13%, N=4) when compared to active control. Timepoint outcomes unable to be pooled showed no difference between groups, including two to three years post-intervention [57,61].

INSERT FIGURE 3

Effect of Peer Support Intervention on Self-efficacy:

Versus usual care (N=6, Figure 4a)

Pooling showed negligible benefit of peer support on self-efficacy relative to usual care in the short-term (SMD: 0.01 [95%CI -0.77, 0.79], P=0.98, I^2=81%, N=2), although peer support resulted in significantly greater effects at both medium-term (SMD: 0.26 [95%CI 0.16, 0.36], P<0.001, I^2=0%, N=4) and long-term (SMD: 0.21 [95%CI 0.07, 0.36], P=0.005, I^2=45%, N=2) timepoints. Studies unable to be pooled had mixed results: one showed no effect at short-term [15] while the other found significantly greater improvements for those receiving peer support (vs usual care) at medium- and long-term follow-up [16].

Versus waitlist control (N=4, Figure 4b)

Only one study assessed self-efficacy at short-term follow-up and showed no significant benefit of peer support over waitlist control [40]. Pooling showed a significant benefit of peer

support on self-efficacy relative to waitlist control at medium-term follow-up (SMD: 0.36 [95%CI 0.20, 0.51], P<0.001, I^2=22%, N=4). No waitlist control studies measured long-term self-efficacy.

Versus active control (N=3, Figure 4c)

One study assessed self-efficacy in the short-term and reported no benefit of peer support relative to an active control [22]. Pooling showed no effect of peer support on self-efficacy relative to active control at medium-term follow-up (MD: -0.09 [95%CI -0.57, 0.40], P=0.73, I^2=27%, N=2). Active controls resulted in significantly greater improvement in self-efficacy at long-term follow-up (MD: -0.41 [95%CI -0.77, -0.05], P=0.03, I^2=0%, N=2). Timepoint outcomes unable to be pooled, including two to three years post-intervention, showed no between group differences [57,61].

INSERT FIGURE 4

Effect of Peer Support Intervention on Function:

Versus usual care (N=9, Figure 5a)

Pooling showed no significant effect of peer support on function relative to usual care in the short- (SMD: -0.04 [95%CI -0.31, 0.23], P=0.77, I^2=57%, N=5) and medium-term (SMD: -0.12 [95%CI -0.25, 0.01], P=0.07, I^2=35%, N=6), but a significantly greater improvement in the long-term (SMD: -0.10 [95%CI -0.19, -0.00], P=0.04, I^2=0%, N=5). One study unable to be pooled found no significant effects [16].

Versus waitlist control (N=7, Figure 5b)

Findings from one study showed no difference in function between peer support and waitlist control in the short-term [21]. Pooling showed no effect of peer support relative to waitlist controls in the medium-term (SMD: -0.10 [95%CI -0.23, 0.04], P=0.16, I^2=23%, N=6).

Versus active control (N=7, Figure 5c)

Pooling showed no effect of peer support on function relative to an active control for short-term (SMD: 0.07 [95%CI -0.17, 0.30], P=0.57, I^2=0%, N=3), medium-term (SMD: -0.10 [95%CI -0.30, 0.10], P=0.32, I^2=0%, N=3), or long-term (SMD: 0.03 [95%CI -0.22, 0.29], P=0.80, I^2=50%, N=4) follow-up. Timepoint outcomes unable to be pooled showed no difference between groups, including two to three years post-intervention [57,61].

INSERT FIGURE 5

Effect of Peer Support Intervention on Health Service Utilisation:

Versus usual care (N=5, Figure 6a)

Pooling showed no effect of peer support relative to usual care on general practitioner (GP) surgery visits in the medium-term (MD: -0.03 [95%CI -0.20, 0.15], P=0.78, I^2=0%, N=2) or long-term (MD: -0.03 [95%CI -0.22, 0.15], P=0.73, I^2=0%, N=2). Findings were mixed in three studies unable to be pooled: one found no difference between peer support and usual care on health service utilisation at medium- and long-term [58]; one showed significantly higher orthopaedic surgeon visits at long-term in the peer support group relative to usual care but no differences in multiple measures at short-term [1]; and the final, a feasibility RCT, showed significant reductions in inpatient stay (in favour of peer support) and GP nurse visits (in favour of usual care) in the short- and medium-term [4].

Versus waitlist control (N=7, Figure 6b)

Pooling showed no difference in physician visits between peer support and waitlist control groups at medium-term follow up (MD: 0.09 [95%CI -0.35, 0.53], P=0.69, I^2=0%, N=3). Of studies unable to be pooled, effects were mixed: three found no significant effect of peer support on health service utilisation at medium-term [11,36,59], while one showed that peer support reduced physical therapy visits relative to waitlist control at medium-term [40].

Versus active control (N=3)

While unable to be pooled, two studies showed no significant effect of peer support on health service utilisation compared to an active controls, spanning medium- and long-term timepoints [51,57,61], while one showed that peer support reduced physician visits at three years [57].

INSERT FIGURE 6

Effect of Peer Support Intervention on Quality of life (QoL):

High heterogeneity precluded pooling.

Versus usual care (N=4)

All four studies comparing peer support to usual care found no significant effect on QoL assessed at short-, medium- and long-term [1,16,44,74,87,88].

Versus waitlist control (N=1)

The sole study using a waitlist control found statistically significant improvements in one measure of QoL in favour of peer support at six weeks, but not sustained at six months [40].

Versus active control (N=2)

One study comparing peer support to an active control found no significant effects on QoL at medium-term [61], while another showed significant improvement in QoL (physical functioning and general health domains) in favour of the active control group at medium- and long-term timepoints [22].

Effect of Peer Support Intervention on Social support:

Pooling was not possible; no comparison/outcome/timepoint pairing had more than one study. Individual studies found no significant effects at any timepoints relative to usual care (via a feasibility RCT) [4], waitlist control [21,78,79], and active control [21,61,79].

Effect of Peer Support Intervention on Self-management:

Heterogeneity in outcome definition precluded pooling.

Versus usual care (N=3)

A feasibility RCT showed a marked increase in self-management knowledge/behaviours for peer support versus usual care at short-term and medium-term [4]. Another revealed a significant effect of peer support for self-management skill acquisition in the short- but not long-term [1], and one study found no significant effects of peer support on health behaviours at medium- and long-term [58].

Versus waitlist control (N=4)

One study found significant improvements in self-management knowledge and practice for peer support relative to waitlist control in the medium-term [54]. Another found no significant findings for self-management knowledge or behaviour in the short-term [21]. The other two studies assessed health behaviours, significantly favouring peer support in cognitive symptom management, relaxation, flexibility exercise, strength exercise [11] and range of motion exercises [53] relative to waitlist control at four months.

Versus active control (N=2)

One study comparing to active control revealed no significant effect of peer support, however subgroup analyses showed greater activation for managing health in those attending over 50% of the intervention [61]. Another found no significant findings for self-management knowledge or behaviour in the short-term [21].

Other outcomes:

All other self-reported clinical outcomes are summarised in Appendix 5.

Sensitivity analyses:

Findings were unchanged with sensitivity analyses (see Appendix 6), except for pooled effects for peer support interventions on long-term self-efficacy relative to active control. When sensitivity analyses were performed using correlation coefficient variation of +0.1, the pooled result for long-term self-efficacy became nonsignificant; showing no significant effects on long-term self-efficacy when compared to active control.

Publication bias:

No pooled analyses included ≥10 studies, preventing formal evaluation of publication bias [39].

DISCUSSION

This is the first meta-analytical systematic review to provide comprehensive synthesis of evidence for peer support interventions in community-dwelling adults with chronic MSK pain. Our meta-analyses suggest that peer support can improve some self-reported clinical outcomes relative to usual care and waitlist controls in the medium- and long-term, but there was no evidence that peer support interventions conferred greater benefit than active interventions. Evidence certainty was low to very low, precluding definitive conclusions.

Beneficial effects (or not) of peer support were closely tied to the comparison condition. Counterintuitively, greater benefit of peer support was present for comparisons to usual care than for comparisons to waitlist (no-treatment) controls. This is surprising given that waitlist control interventions are often criticised for their potential to overestimate treatment effects

[25]. Several explanations may exist. For example, few waitlist control studies could be pooled, resulting in fewer and smaller meta-analyses, potentially limiting the ability to detect (likely) modest effects. Further, peer support interventions in usual care studies used diverse formats (one-to-one or group, face-to-face or online/phone), while waitlist studies did not (mainly group formats). It is possible that non-significant pooled findings for waitlist control comparisons reflect lower efficacy for group-based formats of peer support interventions, although the absence of direct comparison (individual vs group vs waitlist control) precludes conclusion. High heterogeneity in pooled usual care comparisons that used diverse intervention formats (e.g., medium-term pain outcomes) supports the possibility that differences in peer support interventions themselves may contribute to the clinical effect seen. Regardless, that half (n=4/8) of the pooled wait-list control studies were conducted by the same author raises concern about potential bias and unbalanced representation of evidence. Another possible explanation is the lack of comprehensive reporting on if or how waitlist conditions were monitored, creating uncertainty to whether any co-intervention occurred. Participants on the waitlist may have sought other care, perhaps even peer support, positively influencing their clinical outcomes, and thus leading to smaller effect sizes, and non-significant pooled findings. Strictly monitored three-arm clinical trials are needed to definitively ascertain the relationship of peer support relative to usual care and waitlist control, although there remains debate within the field regarding clinical usefulness of waitlist control comparisons [75].

Peer support interventions showed comparable effectiveness to active intervention controls, bar long-term self-efficacy outcomes which favoured active interventions. Such findings are consistent with comparisons in other fields such as cancer [45], diabetes [7], and HIV [14], when a similar educational intervention is provided by either a peer or a health professional. Comparable efficacy to active interventions (and/or evidence of no harm) supports implementation of peer support interventions as an alternative or adjunct to typical clinician-led care. Most active control studies in this review involved self-management programs administered by peer leaders (intervention group) or health professional instructors (active control group), with structural/curricula parity. Perhaps simply engaging in a research intervention such as a program, regardless of who delivers it, is sufficient to confer beneficial effects, resulting in minimal between-group differences. Delivering self-management through peer support may address potential issues that arise when it is offered by healthcare professionals, potentially alleviating feelings of abandonment and loss of treatment agency

(see Stenner et al. [84]). Pooled effects showed wide confidence intervals for both short- and long-term pain outcomes, suggesting evidence uncertainty, and point estimates favoured active control interventions for most outcomes. However, confidence intervals did not include clinically important effects for any outcome at any timepoint. Rigorously designed non-inferiority trials comparing peer support to active interventions with the context of current health systems are warranted.

Beneficial effects of peer support relative to usual care were limited to later follow-up timepoints (medium- and long-term); there was no benefit immediately post-intervention. Such findings were consistent with waitlist control comparisons. Absence of short-term benefits could be predicted by theories of behaviour change [9,65], which suggest that changes in health behaviours are underpinned by complex interactions between the individual and external factors (including social and physical opportunities) incumbent within the environment. Such changes typically take time to occur, with clinical benefits typically paired to sustained lifestyle modifications [13]. Half of the self-efficacy meta-analyses (i.e., confidence in one's capacity to carry out an action) [9] demonstrated a higher degree of confidence in modest effects favouring peer support. Given self-efficacy is an established mediator of the relationship between pain and disability [49], improvements in self-efficacy may be required before changes in behavioural outcomes are observed in the longer-term. Intervention dosage and duration could also play a crucial role. Three-quarters of included studies (n=18/24, 75%) had intervention durations between four and eight weeks. Research, including within the chronic pain realm, indicates that shorter versions of self-management programs are not as effective as longer versions, suggesting that reinforcement, higher dosage and/or time to solidify learnings may be required for greater impact on complex health outcomes [52,70].

This review clearly identified areas with minimal available evidence and for which any conclusion of peer support (non-)effectiveness is pre-emptive. These outcomes include quality of life, social support, and self-management knowledge and behaviour. Further, while available evidence suggests that peer support does not reduce health service utilisation, methodological and contextual heterogeneity was high, and confounding was likely given high prevalence of co-morbidity (or multi-morbidity) in people with chronic MSK pain that may also drive healthcare utilisation [41,86]. Sustaining self-management behaviour and translating these effects into tangible reductions in health service utilisation might be more

challenging or take longer to materialise. Only one study [57] had follow-up longer than 12 months, thus, follow-up durations may also not have been sufficient to capture meaningful changes.

None of the point estimates for clinical outcomes nor their confidence intervals reached established thresholds for clinical importance (see Table 4), although lack of established minimal clinically important difference (MCID) for some self-efficacy scales (e.g., Arthritis Self-Efficacy Scale) limited evaluation. Precise estimates (via narrow confidence intervals) observed for medium- to long-term self-efficacy and functional outcomes relative to usual care support small positive effects from peer support. Modest clinical changes in favour of peer support may not be surprising given peer support in isolation is unlikely to induce large effects, given past work indicating that education and support are most effective when combined with other interventions such as exercise [35,81]. Exploration of whether adding peer support to comprehensive pain management strategies appears warranted. Given the high prevalence and spectrum of chronic MSK pain presentations, there may also be merit in evaluating whether peer support interventions could be effectively integrated within stratified care models, potentially based on individualised pain severity and/or support needs. Quantitative data evaluating the effects of peer support on more global patient benefit is limited, yet studies incorporating qualitative interviews have more comprehensively captured effects on quality of life and social support, highlighting the vast perceived benefits of making supportive interpersonal connections [5,48,63]. Outcomes related to interpersonal enrichment and empowerment (i.e., connection, rapport, hope) appear important for patients but have received minimal exploration [5]. Future research to better understand these potential social and interpersonal advantages are warranted, ensuring the capture of otherwise missed meaningful benefits, especially considering that chronic pain is often isolating [64,82], with social isolation shown to increase pain and disability [42,92]. Involving patients in study design might also ensure improved assessment of meaningful aspects of peer support.

The current evidence holds limitations that merit further exploration. First, two-thirds of included studies focussed on people with arthritis, underscoring an evidence gap for other MSK pain conditions. Second, given the majority of included peer support interventions were structured curricula-based programs, evidence is likely principally limited to informational support attributes (advice, knowledge), with the extent of emotional and/or appraisal support unclear [26,56]. Emotional and appraisal support tend to be fostered in less formal settings,

where sharing experiences, listening, reassurance and encouragement are the focus [26]. For example, meta-analyses primarily featuring Arthritis and Chronic Disease Self-Management Programs warrant careful interpretation as they predominately portray formalised, pre-delineated education for peer support. It is important to consider that the informational/education components of peer support may hold less significance or importance than emotional and appraisal support components. Interventions might benefit from allocating dedicated time to nurture the social and interpersonal dimensions of interactions, separate from the content delivery. Third, few studies (n=4/24, 17%) adopted one-to-one peer support. While those that did demonstrated few significant between-group findings (identifying adherence issues [61], and need for longer-term reinforcement to maintain effects [15,47]), benefits of individualising educational and self-management interventions to the patient have been demonstrated [66], warranting future research on tailored one-to-one peer support. Fourth, few studies (n=4/24, 17%) utilised online or phone delivery modes, likely restricting recruitment of under-served individuals, such as those in significant pain or limited by location. Given recent development of efficacious self-management application of health professional-led interventions via telehealth for chronic pain [76], such a scalable and accessible delivery mode for peer support interventions seems relevant to pursue. Future peer support interventions should also consider greater length and/or dosage to cement learnings. They need to be informed by qualitative findings such as experiences and preferences, while also considering measuring the potential disadvantages and harms from peer support such as misinformation and unhelpful advice [5]. Longer-term follow ups are crucial for exploration of outcomes reliant on prolonged behaviour change, such as health service utilisation. Studies should clarify the type of support peers may offer (informational, appraisal, emotional). Last, consistent measures of self-management and social support would allow pooling in future reviews.

This review had numerous strengths, it: was prospectively registered; adhered to PRISMA guidelines; used a comprehensive search strategy (including grey literature and hand searching) not limited by publication date/language; and only included evidence from randomised, controlled trials. Our review extended that of Cooper et al. (2014) [23] in scale, enabling meta-analysis, and formally defined peer support, addressing previous ambiguities. Finally, focussing solely on chronic MSK pain conditions overcame pitfalls of previous reviews that have encompassed diverse clinical conditions, which has hindered the ability to provide recommendations for specific population groups [32]. The review also had some

weaknesses. Imputation of data was required for meta-analysis compatibility, and the necessity of SMD use limited interpretation of clinically meaningful effects. However, that our sensitivity analyses confirmed findings when imputations were used and that we back translated (where possible) to calculate mean differences for clinical relevance, reduces the impact of these limitations. As occurs with all meta-analyses, it is possible effect sizes are overstated, given pooling can magnify impact of publication bias, if present. We were unable to formally assess publication bias, which is a limitation. Finally, including studies where peer support was part of a multi-component intervention (thus making it challenging to isolate its effect) reduces certainty of peer support being the primary contributor to clinical effect. However, sensitivity analyses excluding these multi-component studies indicated consistent findings, affirming this inclusion did not impact results.

Peer support interventions hold some promise to deliver patient support and education to community-dwelling adults with chronic MSK pain. Available evidence suggests that such interventions lead to small improvements in pain, function, and self-efficacy in the medium- and long-term relative to usual care, although evidence certainty is low-very low and clinical importance unclear. Effects were smaller for comparisons to waitlist controls, and available evidence suggests that peer support interventions lead to comparable effects as active intervention controls, bar self-efficacy which was greater in active controls at long-term. The evidence for pooled health service utilisation was inconclusive, while outcomes related to self-management, quality of life, and social support were varied. Future work should seek to optimise peer support interventions by exploring diverse formats, refining behaviour change targeting, and confirming non-inferiority compared to health professional-led interventions before implementation.

Acknowledgements: MVW is a PhD Candidate supported by a Research Training Program domestic (RTPd) stipend as a postgraduate student at the University of South Australia. FAB is supported by the John Stuart Colville Fellowship via the Arthritis Foundation of South Australia (AFSA). EM is a PhD Candidate supported by a Research Training Program stipend from The Australian Pain Society. TRS is supported by the National Health & Medical Research Council (NHMRC) of Australia Career Development Fellowship (ID1141735). The financial sponsors played no role in study design, collection, analysis, interpretation of data, or manuscript writing/publication. Lorien Delaney reviewed the search strategy.

Conflicts of interest: TRS received payment for lectures on pain and rehabilitation, and receives financial support for travel and accommodation to provide these lectures. TRS receives author royalties for a book on osteoarthritis and pain from OPTP. There are no other conflicts of interest to declare.

REFERENCES

[1] Ackerman IN, Buchbinder R, Osborne RH. Challenges in evaluating an Arthritis Self-Management Program for people with hip and knee osteoarthritis in real-world clinical settings. J Rheumatol 2012;39:1047–1055.

[2] Andersen LN, Juul-Kristensen B, Sørensen TL, Herborg LG, Roessler KK, Søgaard K. Efficacy of tailored physical activity or chronic pain self-management programme on return to work for sick-listed citizens: a 3-month randomised controlled trial. Scand J Public Health 2015;43:694–703.

[3] Andersen LN, Juul-Kristensen B, Sørensen TL, Herborg LG, Roessler KK, Søgaard K. Longer term follow-up of the effects of tailored physical activity or chronic pain self-management programme on return-to-work: A randomized controlled trial. J Rehabil Med 2016;48:887–892.

[4] Anderson AM, Lavender EC, Dusabe-Richards E, Mebrahtu TF, McGowan L, Conaghan PG, Kingsbury SR, Richardson G, Antcliff D, McHugh GA. Peer mentorship to improve self-management of hip and knee osteoarthritis: a randomised feasibility trial. BMJ Open 2021;11:e045389.

[5] Arnott R, Park V, Rhind N, Cooper K. Exploring the experiences and perceptions of participating in a peer-support intervention for adults with chronic non-cancer pain: a qualitative systematic review. JBI Evid Synth 2023;21:1509.

[6] Bair MJ, Matthias MS, Nyland KA, Huffman MA, Stubbs DL, Kroenke K, Damush TM. Barriers and facilitators to chronic pain self-management: a qualitative study of primary care patients with comorbid musculoskeletal pain and depression. Pain Med 2009;10:1280–1290.

[7] Baksi AK, Al-Mrayat M, Hogan D, Whittingstall E, Wilson P, Wex J. Peer advisers compared with specialist health professionals in delivering a training programme on self-management to people with diabetes: a randomized controlled trial. Diabet Med 2008;25:1076–1082.

[8] Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64:401–406.

[9] Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84:191.

[10] Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48:177–187.

[11] Barlow JH, Turner AP, Wright CC. A randomized controlled study of the Arthritis Self-Management Programme in the UK. Health Educ Res 2000;15:665–680.

[12] Behrens F, Koehm M, Schwaneck EC, Schmalzing M, Gnann H, Greger G, Tony H-P, Burkhardt H. Use of a “critical difference” statistical criterion improves the predictive utility of the Health Assessment Questionnaire-Disability Index score in patients with rheumatoid arthritis. BMC Rheumatol 2019;3:51.

[13] Bennell KL, Hinman RS. A review of the clinical evidence for exercise in osteoarthritis of the hip and knee. J Sci Med Sport 2011;14:4–9.

[14] Boucher LM, Liddy C, Mihan A, Kendall C. Peer-led Self-management Interventions and Adherence to Antiretroviral Therapy Among People Living with HIV: A Systematic Review. AIDS Behav 2020;24:998–1022.

[15] Branch VK, Lipsky K, Nieman T, Lipsky PE. Positive impact of an intervention by arthritis patient educators on knowledge and satisfaction of patients in a rheumatology practice. Arthritis Care Res 1999;12:370–375.

[16] Buszewicz M, Rait G, Griffin M, Nazareth I, Patel A, Atkinson A, Barlow J, Haines A. Self management of arthritis in primary care: randomised controlled trial. Bmj 2006;333:879.

[17] Carnes D, Homer KE, Miles CL, Pincus T, Underwood M, Rahman A, Taylor SJ. Effective delivery styles and content for self-management interventions for chronic musculoskeletal pain: a systematic literature review. Clin J Pain 2012;28:344–354.

[18] Chiarotto A, Vanti C, Cedraschi C, Ferrari S, de Lima e Sà Resende F, Ostelo RW, Pillastrini P. Responsiveness and Minimal Important Change of the Pain Self-Efficacy Questionnaire and Short Forms in Patients With Chronic Low Back Pain. J Pain 2016;17:707–718.

[19] Cimmino MA, Ferrone C, Cutolo M. Epidemiology of chronic musculoskeletal pain. Best Pract Res Clin Rheumatol 2011;25:173–183.

[20] Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. What is the Minimum Clinically Important Difference for the WOMAC Index After TKA? Clin Orthop 2018;476:2005–2014.

[21] Cohen JL, van Houten Sauter S, Devellis RF, DeVellis BM. Evaluation of arthritis self-management courses led by laypersons and by professionals. Arthritis Rheum Off J Am Coll Rheumatol 1986;29:388–393.

[22] Coleman S. Effectiveness of Self-Management Programs for People with Osteoarthritis of the Knee [doctoral thesis] Perth, Australia: Curtin University. (2011).

[23] Cooper K, Kirkpatrick P, Wilcock S. The effectiveness of peer support interventions for community-dwelling adults with chronic non-cancer pain: a systematic review. JBI Database Syst Rev Implement Rep 2014;12:319–348.

[24] Cooper K, Schofield P, Klein S, Smith BH, Jehu LM. Exploring peer-mentoring for community dwelling older adults with chronic low back pain: a qualitative study. Physiotherapy 2017;103:138–145.

[25] Cunningham JA, Kypri K, McCambridge J. Exploratory randomized controlled trial evaluating the impact of a waiting list control design. BMC Med Res Methodol 2013;13:150.

[26] Dennis C-L. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003;40:321–332.

[27] Du S, Hu L, Dong J, Xu G, Chen X, Jin S, Zhang H, Yin H. Self-management program for chronic low back pain: A systematic review and meta-analysis. Patient Educ Couns 2017;100:37–49.

[28] Dubé M-O, Langevin P, Roy J-S. Measurement properties of the Pain Self-Efficacy Questionnaire in populations with musculoskeletal disorders: a systematic review. Pain Rep 2021;6:e972.

[29] Egerton T, Lawford BJ, Campbell PK, Plinsinga ML, Spiers L, Mackenzie DA, Graham B, Mills K, Eyles J, Knox G, Metcalf B, Maclachlan LR, Besomi M, Dickson C, Abraham C, Vicenzino B, Hodges PW, Hunter DJ, Bennell KL. Expert-Moderated Peer-to-Peer Online Support Group for People With Knee Osteoarthritis: Mixed Methods Randomized Controlled Pilot and Feasibility Study. JMIR Form Res 2022;6:e32627.

[30] Feldman DE, Carlesso LC, Nahin RL. Management of Patients with a Musculoskeletal Pain Condition that is Likely Chronic: Results from a National Cross Sectional Survey. J Pain 2020;21:869–880.

[31] Forbes R, Mandrusiak A, Russell T, Smith M. Evaluating physiotherapists' practice and perceptions of patient education: A national survey in Australia. Int J Ther Rehabil 2017;24:122–130.

[32] Foster G, Taylor SJ, Eldridge S, Ramsay J, Griffiths CJ. Self-management education programmes by lay leaders for people with chronic conditions. Cochrane Database Syst Rev 2007.

[33] Foster MM, Mitchell G, Haines T, Tweedy S, Cornwell P, Fleming J. Does Enhanced Primary Care enhance primary care? Policy-induced dilemmas for allied health professionals. Med J Aust 2008;188:29–32.

[34] Foster NE, Delitto A. Embedding Psychosocial Perspectives Within Clinical Management of Low Back Pain: Integration of Psychosocially Informed Management Principles Into Physical Therapist Practice—Challenges and Opportunities. Phys Ther 2011;91:790–803.

[35] Goff AJ, De Oliveira Silva D, Merolli M, Bell EC, Crossley KM, Barton CJ. Patient education improves pain and function in people with knee osteoarthritis with better effects when combined with exercise therapy: a systematic review. J Physiother 2021;67:177–189.

[36] Haas M, Groupp E, Muench J, Kraemer D, Brummel-Smith K, Sharma R, Ganger B, Attwood M, Fairweather A. Chronic disease self-management program for low back pain in the elderly. J Manipulative Physiol Ther 2005;28:228–237.

[37] Hägg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003;12:12–20.

[38] Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of

Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res 2011;63:S240–S252.

[39] Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons, 2019.

[40] Hopman-Rock M, Westhoff MH. The effects of a health educational and exercise program for older adults with osteoarthritis for the hip or knee. J Rheumatol 2000;27:1947–1954.

[41] Hudon C, Fortin M, Soubhi H. Chronic musculoskeletal conditions and comorbidities in primary care settings. Can Fam Physician 2008;54:74–75.

[42] Karayannis NV, Baumann I, Sturgeon JA, Melloh M, Mackey SC. The Impact of Social Isolation on Pain Interference: A Longitudinal Study. Ann Behav Med 2019;53:65–74.

[43] Kaya T, Atıcı P, Karatepe AG, Günaydın R. Peer-led education or booklet for knowledge transfer about disease: A randomized-controlled trial with ankylosing spondylitis patients. Arch Rheumatol 2021;36:560.

[44] Kaya T, Goksel Karatepe A, Atici Ozturk P, Gunaydin R. Impact of peer-led group education on the quality of life in patients with ankylosing spondylitis. Int J Rheum Dis 2016;19:184–191.

[45] Klemm P. Effects of Online Support Group Format (Moderated vs Peer-Led) on Depressive Symptoms and Extent of Participation in Women With Breast Cancer. CIN Comput Inform Nurs 2012;30:9.

[46] Kumar D, Neogi T, Peloquin C, Marinko L, Camarinos J, Aoyagi K, Felson DT, Dubreuil M. Delayed timing of physical therapy initiation increases the risk of future opioid use in individuals with knee osteoarthritis: a real-world cohort study. Br J Sports Med 2023;57:958–964.

[47] Laforest S, Nour K, Gignac MA, Gauvin L, Parisien M. The role of social reinforcement in the maintenance of short-term effects after a self-management intervention for frail housebound seniors with arthritis. Can J Aging Rev Can Vieil 2012;31:195–207.

[48] Lavender EC, Anderson AM, Dusabe-Richards E, Antcliff D, Kingsbury SR, Conaghan PG, McHugh GA. Understanding peer mentorship in supporting self-management of hip and knee osteoarthritis: A qualitative study of mentees' perspectives. Musculoskeletal Care 2022;20:180–191.

[49] Lee H, Hübscher M, Moseley GL, Kamper SJ, Traeger AC, Mansell G, McAuley JH. How does pain lead to disability? A systematic review and meta-analysis of mediation studies in people with back and neck pain. PAIN 2015;156:988.

[50] Linton SJ, Hellsing A-L, Larsson I. Bridging the gap: Support groups do not enhance long-term outcome in chronic back pain. Clin J Pain 1997;13:221–228.

[51] Lorig K, Feigenbaum P, Regan C, Ung E, Chastain R, Holman H. A comparison of lay-taught and professional-taught arthritis self-management courses. J Rheumatol 1986;13:763–767.

[52] Lorig K, González VM, Laurent DD, Morgan L, Laris BA. Arthritis self-management program variations: Three studies. Arthritis Rheum 1998;11:448–454.

[53] Lorig K, Gonzalez VM, Ritter P. Community-based Spanish language arthritis education program: a randomized trial. Med Care 1999;37:957–963.

[54] Lorig K, Lubeck D, Kraines RG, Seleznick M, Holman HR. Outcomes of self-help education for patients with arthritis. Arthritis Rheum Off J Am Coll Rheumatol 1985;28:680–685.

[55] Lorig K, Ritter PL, Plant K. A disease-specific self-help program compared with a generalized chronic disease self-help program for arthritis patients. Arthritis Care Res 2005;53:950–957.

[56] Lorig KR, Holman HR. Self-management education: History, definition, outcomes, and mechanisms. Ann Behav Med 2003;26:1–7.

[57] Lorig KR, Ritter PL, Laurent DD, Fries JF. Long-term randomized controlled trials of tailored-print and small-group arthritis self-management interventions. Med Care 2004;42:346–354.

[58] Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Care Res Off J Am Coll Rheumatol 2008;59:1009–1017.

[59] Lorig KR, Sobel DS, Stewart AL, Brown Jr BW, Bandura A, Ritter P, Gonzalez VM, Laurent DD, Holman HR. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999;37:5–14.

[60] Martire LM, Schulz R, Keefe FJ, Rudy TE, Starz TW. Couple-oriented education and support intervention: Effects on individuals with osteoarthritis and their spouses. Rehabil Psychol 2007;52:121.

[61] Matthias MS, Bair MJ, Ofner S, Heisler M, Kukla M, McGuire AB, Adams J, Kempf C, Pierce E, Menen T. Peer support for self-management of chronic pain: The Evaluation of a peer Coach-Led Intervention to improve Pain Symptoms (ECLIPSE) trial. J Gen Intern Med 2020;35:3525–3533.

[62] Matthias MS, Kukla M, McGuire AB, Bair MJ. How Do Patients with Chronic Pain Benefit from a Peer-Supported Pain Self-Management Intervention? A Qualitative Investigation. Pain Med 2016;17:2247–2255.

[63] Matthias MS, Kukla M, McGuire AB, Damush TM, Gill N, Bair MJ. Facilitators and Barriers to Participation in a Peer Support Intervention for Veterans With Chronic Pain. Clin J Pain 2016;32:534–40.

[64] Mellado BH, Falcone AC, Poli-Neto OB, e Silva JCR, Nogueira AA, Candido-dos-Reis FJ. Social isolation in women with endometriosis and chronic pelvic pain. Int J Gynecol Obstet 2016;133:199–201.

[65] Michie S, Van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:1–12.

[66] Moseley GL. Joining forces–combining cognition-targeted motor control training with group or individual pain physiology education: a successful treatment for chronic low back pain. J Man Manip Ther 2003;11:88–94.

[67] Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: The chronic osteoarthritis management initiative of the U.S. bone and joint initiative. Semin Arthritis Rheum 2014;43:701–712.

[68] Ng W, Slater H, Starcevich C, Wright A, Mitchell T, Beales D. Barriers and enablers influencing healthcare professionals' adoption of a biopsychosocial approach to musculoskeletal pain: a systematic review and qualitative evidence synthesis. PAIN 2021;162:2154.

[69] Nicholas M, Vlaeyen JWS, Rief W, Barke A, Aziz Q, Benoliel R, Cohen M, Evers S, Giamberardino MA, Goebel A, Korwisi B, Perrot S, Svensson P, Wang SJ, Treede RD, Iasp Taskforce for the Classification of Chronic Pain. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019;160:28–37.

[70] Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. Diabetes Care 2002;25:1159–1171.

[71] O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, Evans T, Pardo JP, Waters E, White H. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014;67:56–64.

[72] Ørnbjerg LM, Christensen KB, Tennant A, Hetland ML. Validation and assessment of minimally clinically important difference of the unadjusted Health Assessment Questionnaire in a Danish cohort: uncovering ordinal bias. Scand J Rheumatol 2020;49:1–7.

[73] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;10:1–11.

[74] Patel A, Buszewicz M, Beecham J, Griffin M, Rait G, Nazareth I, Atkinson A, Barlow J, Haines A. Economic evaluation of arthritis self management in primary care. Bmj 2009;339.

[75] Patterson B, Boyle MH, Kivlenieks M, Van Ameringen M. The use of waitlists as control conditions in anxiety disorders research. J Psychiatr Res 2016;83:112–120.

[76] Perez J, Niburski K, Stoopler M, Ingelmo P. Telehealth and chronic pain management from rapid adaptation to long-term implementation in pain medicine: A narrative review. Pain Rep 2021;6:e912.

[77] Potter M, Gordon S, Hamer P. The difficult patient in private practice physiotherapy: A qualitative study. Aust J Physiother 2003;49:53–61.

[78] Savelkoul M, de Witte LP. Mutual support groups in rheumatic diseases: Effects and participants' perceptions. Arthritis Care Res 2004;51:605–608.

[79] Savelkoul M, de Witte LP, Candel MJ, van der Tempel H, van den Borne B. Effects of a coping intervention on patients with rheumatic diseases: results of a randomized controlled trial. Arthritis Care Res Off J Am Coll Rheumatol 2001;45:69–76.

[80] Shue SA, McGuire AB, Matthias MS. Facilitators and barriers to implementation of a peer support intervention for patients with chronic pain: a qualitative study. Pain Med 2019;20:1311–1320.

[81] Siddall B, Ram A, Jones MD, Booth J, Perriman D, Summers SJ. Short-term impact of combining pain neuroscience education with exercise for chronic musculoskeletal pain: a systematic review and meta-analysis. PAIN 2022;163:e20.

[82] Smith T. "On their own": social isolation, loneliness and chronic musculoskeletal pain in older adults. Qual Ageing Older Adults 2017;18:87–92.

[83] Spink A, Wagner I, Orrock P. Common reported barriers and facilitators for self-management in adults with chronic musculoskeletal pain: A systematic review of qualitative studies. Musculoskelet Sci Pract 2021;56:102433.

[84] Stenner P, Cross V, McCrum C, McGowan J, Defever E, Lloyd P, Poole R, Moore AP. Self-management of chronic low back pain: Four viewpoints from patients and healthcare providers. Health Psychol Open 2015;2:2055102915615337.

[85] Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, Carpenter J, Rücker G, Harbord RM, Schmid CH. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. Bmj 2011;343:d4002.

[86] Swain S, Sarmanova A, Coupland C, Doherty M, Zhang W. Comorbidities in Osteoarthritis: A Systematic Review and Meta-Analysis of Observational Studies. Arthritis Care Res 2020;72:991–1000.

[87] Taylor SJ, Carnes D, Homer K, Kahan BC, Hounsome N, Eldridge S, Spencer A, Pincus T, Rahman A, Underwood M. Novel three-day, community-based, nonpharmacological group intervention for chronic musculoskeletal pain (COPERS): a randomised clinical trial. PLoS Med 2016;13:e1002040.

[88] Taylor SJ, Carnes D, Homer K, Pincus T, Kahan BC, Hounsome N, Eldridge S, Spencer A, Diaz-Ordaz K, Rahman A. Improving the self-management of chronic pain: COping with persistent Pain, Effectiveness Research in Self-management (COPERS). Programme Grants Appl Res 2016;4.

[89] Tolley JA, Michel MA, Williams AE, Renschler JS. Peer support in the treatment of chronic pain in adolescents: a review of the literature and available resources. Children 2020;7:129.

[90] Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, Bombardier C, Felson D, Hochberg M, van der Heijde D. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005;64:29–33.

[91] Von Korff M, Moore JE, Lorig K, Cherkin DC, Saunders K, González VM, Laurent D, Rutter C, Comite F. A randomized trial of a lay person-led self-management group intervention for back pain patients in primary care. Spine 1998;23:2608–2615.

[92] Wolf LD, Davis MC. Loneliness, daily pain, and perceptions of interpersonal events in adults with fibromyalgia. Health Psychol 2014;33:929–937.

[93] Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003;81:646–656.

[94] World Health Organization. Musculoskeletal health. 2022. Available: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions. Accessed 12 Sep 2023.

FIGURE LEGENDS

Figure 1. PRISMA flowchart of records through the review.

Figure 2. Risk of bias at overall study level.

Figure 3. Pain intensity (0-100 scale) meta-analyses for peer support intervention compared to: a) Usual care; b) Waitlist control; c) Active interventions.
*Peer support was part of a multi-component intervention (effect of peer support alone could not be determined). †Data for peer group intervention without spouses used. ‡Meta-analysis calculated with within group change data. §Multiple follow-up timepoints in same subgrouping; used timepoint that aligned most with other studies in meta-analysis.

Figure 4. Self-efficacy meta-analyses for peer support intervention compared to: a) Usual care; b) Waitlist control; c) Active interventions.
*Data for peer group intervention without spouses used. †Peer support was part of a multi-component intervention (effect of peer support alone could not be determined). ‡Meta-analysis calculated with within group change data. §Multiple follow-up timepoints in same subgrouping; used timepoint that aligned most with other studies in meta-analysis.

Figure 5. Function meta-analyses for peer support intervention compared to: a) Usual care; b) Waitlist control; c) Active interventions.
*Peer support was part of a multi-component intervention (effect of peer support alone could not be determined). †Data for peer group intervention without spouses used. ‡Meta-analysis calculated with within group change data. §Group that received professional-led intervention alone used as active control. **Multiple follow-up timepoints in same subgrouping; used timepoint that aligned most with other studies in meta-analysis.

Figure 6. Health service utilisation meta-analyses for peer support intervention compared to: a) Usual care (GP surgery visits); b) Waitlist control (Physician visits).
*Peer support was part of a multi-component intervention (effect of peer support alone could not be determined). †Meta-analysis calculated with within group change data.

Table 1. Study eligibility criteria using the PICOS (Population, Intervention, Comparison, Outcome, and Study design) framework.

	Inclusion	Exclusion
Population	Community-dwelling adults (≥18 years of age) with chronic MSK pain* (i.e., pain lasting ≥3 months in muscles, bones, joints, or surrounding structures)	Children (<18 years of age);† Participants currently hospitalised or in residential care facilities;† Neurological or cancer pain
Intervention	Peer support interventions, involving peer mentors/volunteers who have the same condition, and have received training as part of the intervention; All formats (1:1, group, face-to-face, online, phone, hybrid, or part of a multi-component intervention)	Intervention delivered by peer mentors/volunteers who have received no training as part of the intervention;‡ Intervention delivered by paraprofessionals§
Comparator	No limit on control type (e.g., usual care, waitlist control, other interventions, sham interventions)	Does not include a control group
Outcomes	Quantitative outcomes including, but not limited to: pain intensity; function/disability; quality of life; self-efficacy; self-management; perceived social support; health service utilisation (any measure); All follow-up timepoints	Only report on qualitative outcomes or only on the outcomes for the peer mentors/volunteers delivering the intervention; No pre- and post-intervention data or group change data
Study design	Randomised controlled trials, randomised cross-over trials, quasi-randomised controlled trials, feasibility randomised controlled trials	Non-randomised trials, case studies, observational designs

*Health conditions not technically termed chronic MSK pain, but for which chronic MSK pain is part of diagnostic criteria (e.g., osteoarthritis) were included. †Studies evaluating peer support interventions in children [89] and for adults hospitalised or living in residential care were excluded as interventions in these populations differ considerably to those provided to community-dwelling adult populations. ‡Termed 'natural-lay-helpers' or 'lay-leaders': informal, untrained support providers (e.g., neighbours) [26]. §'Paraprofessionals': have received very extensive training, non-hierarchical connection is lost [26].

Table 2. Characteristics of included studies.

Author, year, *country*	Intervention Group & Overview	Diagnosis, other criteria	Sample Size, randomized		Age, years mean (SD)	Women, %	Ethnicity	Education	Outcome measures* & follow-up timepoints
Studies with usual care controls only (N=9 studies)									
Ackerman, 2012**[1]** *Australia*	Int: Arthritis Self-Mx Program Con: usual care	Hip and/or knee OA ≥18yo	Int:	n= 58	63.5 (10.8)	62%	69% Australian-born	≤Primary 12%, Yr7-10 47%, Yr11-12 14%, trade 11%, uni 16%	Pain, physical function, quality of life, self-management skills, health service utilisation, disease severity, psychological distress, stiffness *Baseline, 6 weeks, 3 months, 12 months*
			Con:	n= 62	66.6 (10.9)	58%	68% Australian-born	≤Primary 12%, Yr7-10 45%, Yr11-12 17%, trade 18%, uni 8%	
Anderson, 2021**[4]**** *UK*	Int: one-to-one peer mentorship sessions Con: usual care	Hip and/or knee OA ≥55yo	Int:	n= 25	70.0 (8.6)	58%	87.5% white	62.5% further education	Pain, function, self-efficacy, self-management, perceived social support, healthcare & community resource use, health status, anxiety & depression, stiffness *Baseline, 8 weeks, 6 months*
			Con:	n= 25	69.3 (8.1)	84%	96% white	60% further education	
Branch, 1999**[15]** *US*	Int: Arthritis Patient Educator Con: usual care	OA, rheumatoid arthritis or fibromyalgia	Int:	n= 47	(No demographic data reported)				Arthritis impact (health status), self-efficacy, knowledge *Baseline, 8 weeks*
			Con:	n= 61					
Buszewicz, 2006**[16]** Patel, 2009**[74]** *UK*	Int: Arthritis Self-Mx Program Con: usual care	Hip and/or knee OA ≥50yo	Int:	n= 406	68.4 (8.2)	63%	n=390: 388 white, 0 Black African, 2 Black Caribbean	n=388: 28% higher education	Buszewicz 2006: pain, quality of life, physical functioning, self-efficacy, anxiety & depression; Patel 2009: quality of life, resource use, mental health, physical health, health state *Baseline, 4 months, 12 months*
			Con:	n= 406	68.7 (8.6)	63%	n=385: 382 white, 1 Black African, 2 Black Caribbean	n=382: 27% higher education	
Kaya, 2016**[44]** Kaya, 2021**[43]** *Turkey*	Int: peer education group + booklet Con: usual care	Ankylosing spondylitis 18-75yo	Int:	n= 40	43.1 (9.1)	22%		Duration of education: Median 7 (5-19)	Kaya 2016: functional status, activity status, health status, quality of life, depression; Kaya 2021: knowledge *Baseline, 4 weeks, 6 months*
			Con:	n= 40	40.9 (9.3)	14%		Median 8 (0-15)	
Lorig, 2008**[58]** *US*	Int: Internet-based Arthritis Self-Mx program Con: usual care	OA, rheumatoid arthritis or fibromyalgia	Int:	n= 433	52.2 (10.9)	89.8%	90.9% white	15.6 (3.09) mean education years	Pain, disability, self-efficacy, health-related behaviors, healthcare utilisation, global
			Con:	n= 422	52.5 (12.2)	90.5%	93.7% white	15.7 (3.11)	

		≥18yo							health, fatigue, health distress, activity limitations *Baseline, 6 months, 12 months*
Martire, 2007**[60]** *US*	Int A: patient education & support Int B: couples-orientated education & support Con: usual care	Hip and/or knee OA ≥50yo Married	Int A:	n= 89	68.0 (8.0)	72%		14.6 (1.7) mean education years	Pain, physical function, self-efficacy, marital satisfaction, stiffness, depressive symptoms *Baseline, 6 weeks, 6 months*
			Int B:	n= 54	68.4 (7.5)	72%		14.2 (1.5)	
			Con:	n= 99	69.2 (7.2)	74%		14.3 (1.6)	
Taylor, 2016a**[87]** Taylor, 2016b**[88]** *UK*	Int: Self-Mx group course Con: usual care	Chronic MSK pain ≥18yo	Int:	n= 403	60.3 (13.5)	67%	81% white, 13% black, 3% Asian, 3% mixed/other	Age formal education ended: 56% ≤16yo, 43% ≥20yo, 1% other	Taylor 2016a: Pain, disability, self-efficacy, social integration, health utility (quality of life), depression & anxiety, pain acceptance & coping, global health, defined daily doses; Taylor 2016b: as above + healthcare utilisation *Baseline, 6 months, 12 months*
			Con:	n= 300	59.4 (13.8)	67%	80% white, 12% black, 7% Asian, <1% mixed/other	52% ≤16yo, 45% ≥20yo, 3% other	
Von Korff, 1998**[91]** *US*	Int: lay-led Self-Mx Group Con: usual care	Back pain, strain, disc disorder or sciatica 25-70yo	Int:	n= 129	49.4 (11.7)	68.2%	91.4% white, 8.6% non-white	48.1% college grad, 39.5% some college, 12.4% <Yr12	Pain, pain interference, impairment & limitation, self-care attitudes, back pain worries, mental health *Baseline, 3 months, 6 months, 12 months*
			Con:	n= 126	50.3 (10.9)	56.4%	79.7% white, 20.3% non-white	45.6% college grad, 37.6% some college, 16.8% <Yr12	
Studies with waitlist controls only (N=6 studies)									
Barlow, 2000**[11]** *UK*	Int: Arthritis Self-Mx Program Con: wait-list control	Arthritis ≥18yo	Int:	n= 344	57.3 (13.2)	85%	98% Caucasian	52% some formal education qualification	Pain, physical functioning, self-efficacy, use of cognitive symptom management, health behaviors, visits to GP, health status, fatigue, psychological wellbeing *Baseline, 4 months (12 months 1 group only)*
			Con:	n= 258	59.1 (12.3)	83%	94% Caucasian	52% some formal education qualification	
Haas, 2005**[36]** *US*	Int: Chronic Disease Self-Mx Program Con: wait-list control	Chronic low back pain ≥60yo	Int:	n= 60	78.6 (7.5)	81.6%	18.4% African American, 81.6% White	High school grad 93.2%, College grad 18.2%	Pain, functional disability, self-efficacy, health care utilisation, general health, emotional wellbeing, self-care attitudes, pain days, disability days *Baseline, 6 months*
			Con:	n= 60	75.5 (7.5)	87.8%	10.2% African American, 89.8% White	High school grad 97.5%, College grad 30.0%	
Hopman-Rock, 2000**[40]**	Int: Living with OA Program Con: wait-list control	Hip and/or knee OA 55-75yo	Int:	n= 60	65.4 (5.3)	78%		17% primary, 54% secondary, 27% college/uni	Pain, mobility, quality of life, self-efficacy, health care utilisation,

Netherlands			Con:	n= 60	65.2 (5.7)	88%		26% primary, 45% secondary, 20% college/uni	body mass index, knowledge, observed activity restrictions *Baseline, 6 weeks, 6 months*
Lorig, 1985**[54]** *US*	Int: Arthritis Self-Mx Program Con: wait-list control	Arthritis	Int:	n= 134	67.4 (11.84)†	83%†			Pain, physical disability, self-management activities, health service utilisation, knowledge, locus of control *Baseline, 4 months*
			Con:	n= 65					
Lorig, 1999a**[59]** *US*	Int: Chronic Disease Self-Mx Program Con: wait-list control	Heart disease, or lung disease, or stroke or arthritis‡ ≥40yo	Int:	n= 664‡ (86 arthritis)	65.6 (range 40-90)‡	65%‡	91.4% white‡	Mean years: 27% ≤12, 28% 13-15, 16% 16, 29% >16‡	Pain/physical discomfort, disability, health service utilisation, self-rated health, psychological wellbeing, energy/fatigue, health distress, health behaviours, social/role activity limitations, shortness of breath *Baseline, 6 months*
			Con:	n= 476‡ (62 arthritis)	65.0 (range 40-89)‡	64%‡	88.8% white‡	27% ≤12, 25% 13-15, 21% 16, 27% >16‡	
Lorig, 1999b**[53]** *US*	Int: Spanish Arthritis Self-Mx Program Con: wait-list control	Arthritis Spanish-speaking	Int:	n= 219	62.5 (range 29-93)	85%	51% Mexico, 31% Central America, 15% Sth America, 3% Caribbean	8.1 mean education years	Pain, disability, self-efficacy, self-management behaviour, visits to physician, self-rated health, depression, medication use *Baseline, 4 months*
			Con:	n= 112	62.5 (range 18-87)	81%	49% Mexico, 33% Central America, 14% Sth Amercia, 4% Caribbean	8.1 mean education years	
Studies with active controls only (N=3 studies)									
Coleman, 2011**[22]** *Australia*	Int: lay-led Arthritis Self-Mx Program Con: active (professional-led Self-Mx program)	Knee OA ≥18yo	Int:	n= 90	66.3 (9.84)	68%			Pain, physical function, quality of life, self-efficacy, global health, step test, single leg balance, TUG *Baseline, 8 weeks, 6 months, 12 months*
			Con:	n= 90	67.6 (8.23)	67%			
Lorig, 2004**[57]** *US*	Int: Arthritis Self-Mx Program Con: active (mail delivered self-Mx program)	Arthritis (OA or RA) Adults	Int:	n= 161	65.2 (range 22-90)	75%	90% white	14.7 mean education years (range 3-22)	Pain, disability, self-efficacy, healthcare utilisation, depression, global severity of arthritis, role function *Baseline, 12 months, 24 months, 36 months*
			Con:	n= 180					
Matthias, 2020**[61]** *US*	Int: 'ECLIPSE' one-to-one peer support Con: active (self-Mx class)	Chronic MSK pain Veterans	Int:	n= 120	55.4 (12.6)	20.2%	63.9% white, 1.7% hispanic	76.3% >high school	Pain, physical functioning, quality of life, self-efficacy, self-management, perceived social support, healthcare utilisation,
			Con:	n= 95	58.6 (13.3)	17%	58.5% white, 4.3% hispanic	77.7% >high school	

									general health perceptions, anxiety & depression, pain coping *Baseline, 6 months, 9 months*
Studies with multiple control groups (N=6 studies)									
Andersen, 2015**[2]** Anderson, 2016**[3]** *Denmark*	Int: Chronic Disease Self-Mx Program Con A: usual care♦ Con B: active (Tailored Physical Activity Group)♦	Sick-listed citizens due to back/upper body pain	Int:	n= 47	44.3 (10.8)	60%		n=44, education: 21% no, 45% low, 28% medium-high	Pain, work ability, body mass index, return to work % & days taken to return, kinesiophobia, aerobic capacity, hand grip strength *Baseline, 3 months* (Anderson 2015)*, 11 months* (Anderson 2016)
			Con A:	n= 47	45.8 (10.8)	57%		n=47, education: 15% no, 51% low, 34% medium-high	
			Con B:	n= 47	45.6 (10.0)	50%		n=43, education: 17% no, 54% low, 24% medium-high	
Cohen, 1986**[21]** *US*	Int: lay-led Arthritis Self-Mx Course Con A: wait-list control♦ Con B: active (professional-led Arthritis Self-Mx Course)♦	Arthritis	Int:	n= 32	65.5	78%	95% white	16.2 mean years	Pain, functional disability, knowledge, self-management behaviors, perceived affective & instrumental support, depression *Baseline, 6 weeks, 3 months*
			Con A:	n= 36					
			Con B:	n= 28					
Laforest, 2012**[47]** *Canada*	Int: social reinforcement with self-Mx intervention Con A: active (Self-Mx intervention only)♦ Con B: wait-list control	OA or rheumatoid arthritis ≥50yo Housebound	Int:	n= 29	77.5 (10.3)	96.6%		9.21 (4.2) mean education years	Functional limitations, coping, helplessness *Baseline, 2 months, 10 months*
			Con A:	n= 36	76.6 (11.1)	86.1%		10.49 (4.4)	
			Con B:	n=48	78.7 (10.2)	89.6%		8.14 (3.4)	
Linton, 1997**[50]** *Sweden*	Int: lay-led educational support group Con A: usual care♦ Con B: active (professional-led support group)♦	Chronic MSK pain Accumulated sick leave 2-24 wks/past year 18-60yo	Int:	n= 39	50 (9.9)	74%			Pain, function, coping strategies, health status, sick leave, pain beliefs & attitudes, overall outcome evaluation *Baseline, 12 months*
			Con A:	n= 25	53 (9.6)	68%			
			Con B:	n= 39	50 (9.6)	77%			
Lorig, 1986**[51]** *US*	Int: lay-led Arthritis Self-Mx Course Con A: wait-list control♦ Con B: active (professional-led	Arthritis	Int:	n= 34	69.8 (8.8)	72%		13.4 (2.8) mean education years	Pain, disability, self-management knowledge, self-management behaviours, visits to physician, exercise frequency *Baseline, 4 months*
			Con A:	n= 32	61.6 (12.2)	79%		15.0 (4.3)	
			Con B:	n= 34	62.1 (14.1)	69%		13.9 (3.9)	

	Arthritis Self-Mx Course)♦							
Savelkoul, 2001**[79]** Savelkoul, 2004**[78]**§ *Netherlands*	Int: mutual support group Con A: wait-list control Con B: active (coping intervention group)	Chronic rheumatic disorder >1 yr 35-65yo	Int:	n= 56	51.1 (8.91)	58.9%	Level of education: 40.4% low, 46.8% medium, 12.8% high	Savelkoul 2001: Social interactions, functional health status, mobility, action directed coping, loneliness, life satisfaction, coping by seeking social support; Savelkoul 2004: Social network size, social skills, functional health status, loneliness, life satisfaction *Baseline, 6 months*
			Con A:	n= 56	50.5 (8.65)	67.9%	51.1% low, 28.9% medium, 20% high	
			Con B:	n= 56	52.5 (8.31)	76.8%	54.2% low, 35.4% medium, 10.4% high	

♦Control group/s used in meta-analyses. *Feasibility measures not listed under outcome measures. †Reports on demographic data for the ASMP intervention group only. ‡Demographic data presented for entire study, of which only some participants had arthritis. §Paper reported on mutual support group versus wait list control group only. **Study design: feasibility randomised controlled trial. Int = intervention group; Con = control group; Mx = management.

Table 3. Intervention and control condition characteristics of included studies.

Author, year	Intervention & control conditions overview	Format	Providers	Duration	Adherence or attendance rates	Content
Studies with usual care controls only						
Ackerman, 2012**[1]**	Int: Arthritis Self-Mx Program	Group, F2F	1 peer leader & 1 health professional*	6 weeks (1x 2.5hr session per week)	47% attended all sessions; Median 5 sessions attended	Topics: pain, fatigue, physical activity, managing emotions, health-related problem-solving, communication with doctors
	Con: usual care (education book)	Paper	-	-		Arthritis self-help book
Anderson, 2021**[4]**‡	Int: one-to-one peer mentorship sessions	1:1, F2F	Trained peer mentors	8 weeks (1x 1hr session per week)	83% completion rate; Mean 5.79 sessions received	Topics: behaviour change techniques, learning about OA, goal setting, pacing, muscle strengthening, connecting with others, optional topics
	Con: usual care (information resources)	1:1, F2F	Researcher	1 session		Arthritis information resources discussed with researcher
Branch, 1999**[15]**	Int: Arthritis Patient Educator (before/after usual rheumatology appointment)	1:1, F2F & phone	Trained person with arthritis	3 sessions (pre-appointment contact, F2F session post-appointment, phone call 1 week later)	57% completed entire protocol	Peer support, education, Arthritis Foundation pamphlets, possible referral to the social worker
	Con: usual care (rheumatology appointment)	1:1, F2F	Rheumatologist	1 appointment		Usual rheumatologist appointment
Buszewicz, 2006**[16]** Patel, 2009**[74]**	Int: Arthritis Self-Mx Program group	Group, F2F	Trained facilitator pair, at least one had arthritis	6 weeks (1x 2-2.5hr session per week)		Topics: exercising, medication, communicating, nutrition, sleep, informed treatment, problem-solving
	Con: usual care (education booklet)	Paper	-	-		Arthritis education booklet
Kaya, 2016**[44]** Kaya, 2021**[43]**	Int: Peer education group + booklet	Group, F2F	Trained peer educators	4 weeks (1x 1hr session per week)	32.5% attended no sessions	Topics: etiology, clinical findings, complications, treatment, exercise importance, joint protection
	Con: usual care (education booklet)	Paper	-	-		Booklet constituted of the same topics as the education sessions
Lorig, 2008**[58]**	Int: Internet-based Arthritis Self-Mx program	Self-directed & group, online	Trained peer moderators	6 weeks (3 logins per week, 1-2hrs per login)	Mean 31.6 log-ins; 6% never logged in	Topics: exercise programs, cognitive symptom management, managing emotions, medications, communication, healthy eating, fatigue, action planning Web-based bulletin board discussion, tools/logs, Arthritis Helpbook
	Con: usual care	-	-	-		Continued with their usual care
Martire, 2007**[60]**	Int A: lay-led patient education & support	Group, F2F	Trained lay leader pair	6 weeks (1x 1hr session per week)	76% attended ≥1 session	Topics: etiology, treatments, self-management, exercise benefits, communication, coping

	Int B: couples-orientated education & support	Group, F2F, *with spouses*	Trained lay leader pair	6 weeks (1x 1hr session per week)	72% attended ≥1 session	Topics framed to couple: etiology, treatments, self-management, exercise benefits, communication, coping
	Con: usual care (normal medical regime & rheumatology visits)	-	Included rheumatologist	-		Usual medical regimes and rheumatologist appointments
Taylor, 2016a**[87]** Taylor, 2016b**[88]**	Int: 'COping with persistent Pain, Effectiveness Research into Self-management' group course	Group, F2F	Trained facilitator pair (1 lay, 1 health professional)*	6 weeks (1x 2hr session per week)	85% sessions attended; 17% attended no sessions	Topics: exercise, relaxation, joint protection, heat, massage, medications, diet, communication, solving problems
	Con: usual care (booklet & relaxation CD)	Self-directed	-	3 weeks minimum (relaxation practiced per day)		Relaxation CD (practice each day), Pain Toolkit booklet
Von Korff, 1998**[91]**	Int: lay-led Self-Mx Group	Group, F2F	Trained lay leaders, one had arthritis	4 weeks (1x 2hr session per week)	68% attended ≥3 sessions; 11% attended no sessions	Topics: pacing, exercise, posture, mechanics, self-talk, flare-ups, self-care strategies, goals
	Con: usual care (plus education book)	-	-	-		'Your Aching Back' book
Studies with waitlist controls only						
Barlow, 2000**[11]**	Int: Arthritis Self-Mx Program group	Group, F2F	Trained lay leaders	6 weeks (1x 2hr session per week)		Topics: information about arthritis, self-management techniques, exercise, cognitive self-management, nutrition, goal setting
	Con: 4mo wait-list control	-	-	-		-
Haas, 2005**[36]**	Int: Chronic Disease Self-Mx Program	Group, F2F	Trained lay leaders	6 weeks (1x 2.5hr session per week)	68% attended ≥3 classes; 16% attended all classes	Topics: self-management principles, care-seeking options, community resources, exercise, relaxation, nutrition, medication, skills building, goal setting, action plans
	Con: 6mo wait-list control	-	-	-		-
Hopman-Rock, 2000**[40]**	Int: Living with OA Program	Group, F2F	Education by trained peer, exercise by physiotherapist*	6 weeks (1x 2hr session per week)	67% attended all sessions	Topics: pathophysiology, lifestyle, physical activity, pain management, weight, diet Exercise program with physiotherapist
	Con: 6mo wait-list control	-	-	-		-
Lorig, 1985**[54]**	Int: Arthritis Self-Mx Program	Group, F2F	Trained lay leader pair	16 weeks (6 sessions total)		Topics: arthritis information, medication use, exercise, relaxation, joint protection, nutrition
	Con: 4mo wait-list control	-	-	-		-
Lorig, 1999a**[59]**	Int: lay-led Chronic Disease Self-Mx Program	Group, F2F	Trained lay leader pair	7 weeks (1x 2.5hr session per week)	Mean 5.5 sessions attended	Topics: education about arthritis, exercises, relaxation, medications, nutrition, joint protection
	Con: 6mo wait-list control	-	-	-		-
Lorig, 1999b**[53]**	Int: lay-led Spanish Arthritis Self-Mx Program	Group, F2F	Trained lay leaders	6 weeks (1x 2hr session per week)		Exercise in class, education, how to access care, Living with Arthritis book, audio exercise &

						relaxation tapes, illustrated exercise book (cultural/language adaptation)
	Con: 4mo wait-list control	-	-	-		-
Studies with active controls only						
Coleman, 2011**[22]**	Int: lay-led Arthritis Self-Mx program	Group, F2F	Trained lay leaders	6 weeks (1x 2.5hr session per week)		Topics: pain, fatigue, exercises, use of medications, communicating, eating healthy, informed treatment decisions, problem solving, sleep
	Con: active (OAK professional-led Self-Mx program)	Group, F2F	Health professionals	6 weeks (1x 2.5hr session per week)		Detailed exercise instruction, pain management, evidence-based information, pathophysiology, self-management
Lorig, 2004**[57]**	Int: Arthritis Self-Mx Program	Group, F2F	Trained lay leaders	6 weeks (1x 2hr session per week)	Mean 4.6 sessions attended	Topics: knowledge, nutrition, medication, cognitive restructuring techniques, physical activity advice, problem-solving, improving communication
	Con: active (mail delivered tailored self-Mx program)	Mailed, paper	Planned via algorithm	12-18 months (new plan sent quarterly)		Content consistent with intervention course but more tailored to person based on questionnaire responses
Matthias, 2020**[61]**	Int: 'ECLIPSE' one-to-one peer support	1:1, F2F or phone	Trained peer coaches	6 months (2x sessions per month)	13.1% attended ≥12 sessions; 64% attended ≤5 sessions; 8% attended no sessions	Topics: pain self-management, relaxation, pacing, cognitive behavioural techniques, skills, self-care
	Con: active (self-Mx class)	Group, F2F	Facilitators	1x 2-hour session		Topics: chronic pain basics, relaxation skills, activity pacing
Studies with multiple control groups						
Andersen, 2015**[2]** Anderson, 2016**[3]**	Int: Chronic Disease Self-Mx Program	Group, F2F	Trained peer facilitators	6 weeks (1x 2.5hr session per week)	50% adherence	Techniques to deal with fatigue, use of medications, mutual support, encouragement to stay active
	Con A: usual care (health guidance only)♦	1:1, F2F	Health supervisors	1x 1.5hr session		Goal-oriented health plan with guidance and support
	Con B: active (Tailored Physical Activity Group)♦	Group, F2F	Physiotherapists	10 weeks (3x 50min sessions per week)		5-minute warm-up followed by aerobic fitness training, with progressions
Cohen, 1986**[21]**	Int: lay-led Arthritis Self-Mx Course	Group, F2F	Trained lay leaders	6 weeks (1x 2hr session per week)		Topics: exercise, relaxation, joint protection, heat, massage, medications, diet
	Con A: wait-list control♦	-	-	-		-
	Con B: active (professional-led Arthritis Self-Mx Course)♦	Group, F2F	Health professionals	6 weeks (1x 2hr session per week)		Exercises, equipment use, knowledge

Laforest, 2012**[47]**	Int: social reinforcement with self-Mx intervention	1:1, phone	Trained peers	6 months (8x 15min calls: bi-monthly for 2mo, monthly for 4mo)		Positive feedback, stimulated reflection, problem-solving activities, action plan, course topics
	Con A: active (self-Mx intervention only)♦	1:1, F2F	Healthcare practitioners	6 weeks (1x 1hr session per week)		Pain management, exercise, relaxation, coping, support, goals and action plans, review
	Con B: wait-list control	-	-	-		-
Linton, 1997**[50]**	Int: lay-led educational support group	Group, F2F	Trained lay person	1 year (15x 3hr sessions over span of year)	Mean 8.9 sessions received	Topics: pain physiology, attitudes, strategies for pain, exercise
	Con A: usual care (medical and allied health care)♦	F2F	Normal GP & allied health contact	1 year		Could include GP contact, analgesics, physical therapy/specialist/rehab if needed
	Con B: active (professional-led support group)♦	Group, F2F	Trained professionals	1 year (15x 3hr sessions over span of year)		Topics: cognitive behavioural therapy, emotional support, controlling pain, activity limitations, family/social/workplace issues, medications, stress management, relaxation, exercises
Lorig, 1986**[51]**	Int: lay-led Arthritis Self-Mx Course	Group, F2F	Trained lay leader pair	6 weeks (1x 2hr session per week)	Mean 5 sessions attended	Topics: education about arthritis, exercises, relaxation, medications, nutrition, problem solving, joint protection, therapies, communication
	Con A: wait-list control♦	-	-	-		-
	Con B: active (professional-led Arthritis Self-Mx Course)♦	Group, F2F	Rheumatologist & physical therapist	6 weeks (1x 2hr session per week)		Exercises, equipment use, knowledge
Savelkoul, 2001**[79]** Savelkoul, 2004**[78]**†	Int: mutual support group	Group, F2F	Trained patient pair	13 weeks (1x 2hour session per week for 8 weeks; next session 2 weeks later; next session 3 weeks later)	Mean 6.4 sessions attended	Exchange information, experiences, feelings, emotions alongside topics
	Con A: wait-list control	-	-	-		-
	Con B: active (coping intervention group)	Group, F2F	Behavioral therapist & nurse or social worker	13 weeks (1x 2hour session per week for 8 weeks; next session 2 weeks later; next session 3 weeks later)		Aimed at increasing active-directed coping and coping by seeking social support

♦Control group/s used in meta-analyses. *Intervention provided part by peer and part by non-peer (so unable to determine effect of peer support alone). †Paper reported on mutual support group versus wait list control group only. ‡Study design: feasibility randomised controlled trial. Int = intervention group; Con = control group; Mx = management; F2F = face-to-face; 1:1 = one-to-one.

Table 4. Summary of meta-analysis results.

Control	*Timepoint*	*Pooled effect*	*Effect size [95%CI], P-value*	*GRADE certainty of evidence*	*Interpretation*	*Met MCID (yes/no)*	*Back-transformation value*
Figure 3. Pain intensity meta-analyses.							
a. Usual Care	Short-Term	Not significant	MD: -1.58 [-4.83, 1.66], P=0.34	Low		No*	*No back-transformation required as MDs used, see Figure 3.*
	Medium-Term	SIGNIFICANT	**MD: -3.48 [-6.61, -0.35], P=0.03**	Very low	Favours peer support	No*	
	Long-Term	SIGNIFICANT	**MD: -1.97 [-3.53, -0.42], P=0.01**	Low	Favours peer support	No*	
b. Waitlist	Short-Term	Not significant	MD: 2.00 [-5.12, 9.12], P=0.58	Very low		No*	
	Medium-Term	Not significant	MD: -2.90 [-6.62, 0.81], P=0.13	Very low		No*	
c. Active Control	Short-Term	Not significant	MD: 4.98 [-0.08, 10.04], P=0.05	Very low		No*	
	Medium-Term	Not significant	MD: 1.90 [-1.79, 5.59], P=0.31	Very low		No*	
	Long-Term	Not significant	MD: 2.94 [-0.01, 5.90], P=0.05	Low		No*	
Figure 4. Self-efficacy meta-analyses.							
a. Usual Care	Short-Term	Not significant	SMD: 0.01 [-0.77, 0.79], P=0.98	Very low		N/A†	0.02 (95%CI -1.39 to 1.42)†
	Medium-Term	SIGNIFICANT	**SMD: 0.26 [0.16, 0.36], P<0.001**	Low	Favours peer support	No‡	3.63 (95%CI 2.23 to 5.16)‡
	Long-Term	SIGNIFICANT	**SMD: 0.21 [0.07, 0.36], P=0.005**	Very low	Favours peer support	No‡	2.93 (95%CI 0.98 to 5.02)‡
b. Waitlist	Medium-Term	SIGNIFICANT	**SMD: 0.36 [0.20, 0.51], P<0.001**	Low	Favours peer support	N/A†	0.85 (95%CI 0.47 to 1.20)†
c. Active Control	Medium-Term	Not significant	MD: -0.09 [-0.57, 0.40], P=0.73	Very low		N/A†	*No back-transformation required as MDs used, see Figure 4.*
	Long-Term	SIGNIFICANT	**MD: -0.41 [-0.77, -0.05], P=0.03**	Low	Favours control	N/A†	
Figure 5. Function meta-analyses.							
a. Usual Care	Short-Term	Not significant	SMD: -0.04 [-0.31, 0.23], P=0.77	Very low		No§	-0.82 (95%CI -6.37 to 4.73)§
	Medium-Term	Not significant	SMD: -0.12 [-0.25, 0.01], P=0.07	Low		No§	-1.34 (95%CI -2.80 to 0.11)§

	Long-Term	SIGNIFICANT	**SMD: -0.10 [-0.19, -0.00], P=0.04**	Very low	Favours peer support	No§	-2.06 (95%CI -4.11 to -0.21)§
b. Waitlist	Medium-Term	Not significant	SMD: -0.10 [-0.23, 0.04], P=0.16	Very low		No**	-0.06 (95%CI -0.14 to 0.02)**
c. Active Control	Short-Term	Not significant	SMD: 0.07 [-0.17, 0.30], P=0.57	Very low		No§	0.84 (95%CI -2.04 to 3.6)§
	Medium-Term	Not significant	SMD: -0.10 [-0.30, 0.10], P=0.32	Very low		No§	-1.20 (95%CI -3.6 to 1.2)§
	Long-Term	Not significant	SMD: 0.03 [-0.22, 0.29], P=0.80	Low		No§	0.36 (95%CI -2.64 to 3.48)§

Figure 6. Health service utilisation meta-analyses.

a. Usual Care	Medium-Term	Not significant	MD: -0.03 [-0.20, 0.15], P=0.78	Very low		N/A	*No back-transformation required as MDs used, see Figure 6.*
	Long-Term	Not significant	MD: -0.03 [-0.22, 0.15], P=0.73	Very low		N/A	
b. Waitlist	Medium-Term	Not significant	MD: 0.09 [-0.35, 0.53], P=0.69	Very low		N/A	

*MCID for pain intensity is 11-19 points on 0-100 scale (for various MSK pain populations) [37,38,90]. †No interpretability data available for self-efficacy on 1-10 Arthritis Self-Efficacy Scale (ASES). ‡MCID for self-efficacy is 5.5-8.5 points on 0-60 Pain Self-Efficacy Questionnaire (PSEQ) [18,28]. §MCID for function is 9 points on 0-68 function subscale of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [20]. **MCID for function is 0.48-0.68 on 0-3 function subscale of Health Assessment Questionnaire (HAQ) [12,72].

Figure 1 rev

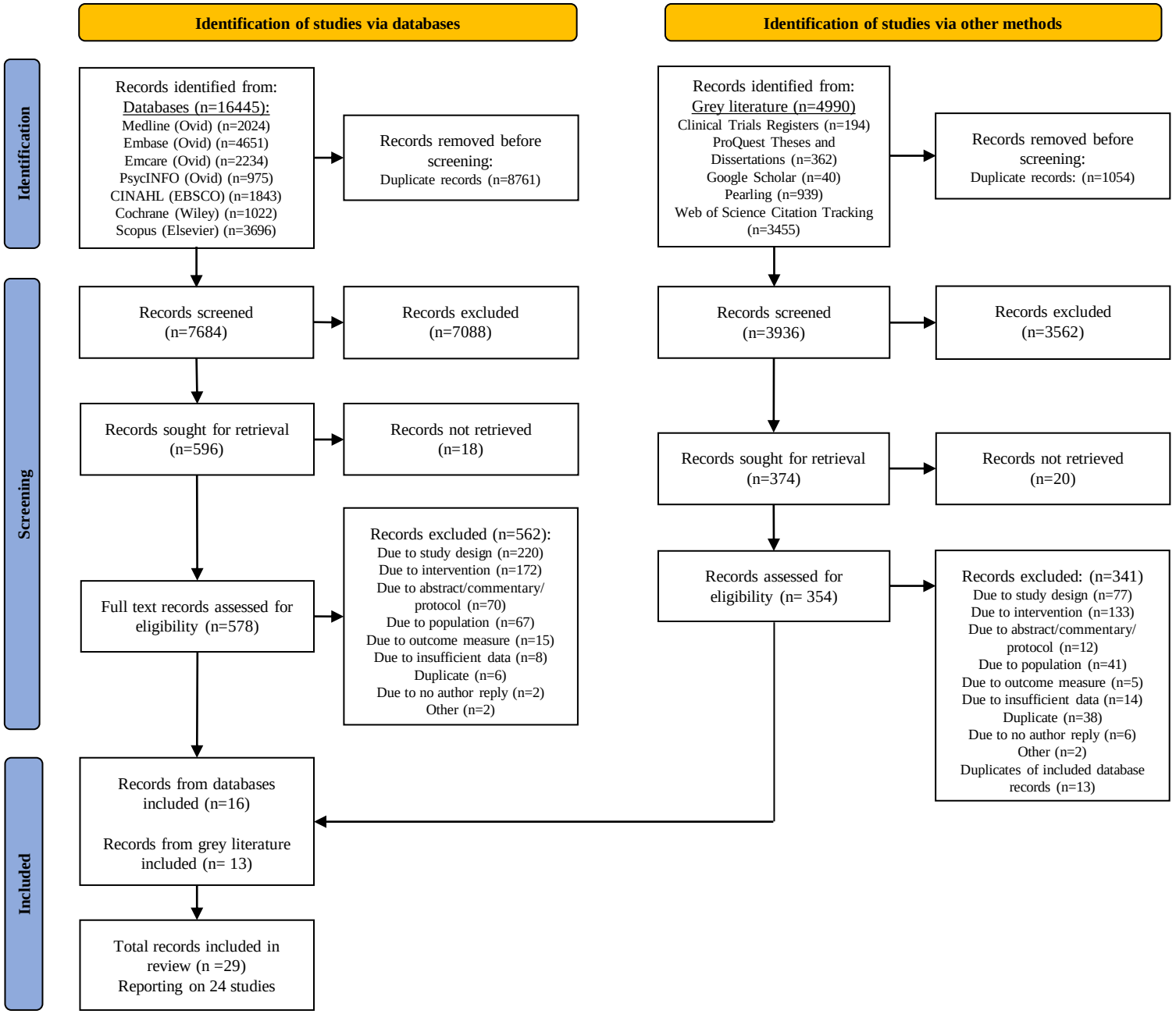

Identification of studies via databases
Identification of studies via other methods
Identification
Screening
Included
Records identified from:
Databases (n=16445):
Medline (Ovid) (n=2024)
Embase (Ovid) (n=4651)
Emcare (Ovid) (n=2234)
PsycINFO (Ovid) (n=975)
CINAHL (EBSCO) (n=1843)
Cochrane (Wiley) (n=1022)
Scopus (Elsevier) (n=3696)
Records removed before screening:
Duplicate records (n=8761)
Records identified from:
Grey literature (n=4990)
Clinical Trials Registers (n=194)
ProQuest Theses and Dissertations (n=362)
Google Scholar (n=40)
Pearling (n=939)
Web of Science Citation Tracking (n=3455)
Records removed before screening:
Duplicate records: (n=1054)
Records screened (n=7684)
Records excluded (n=7088)
Records screened (n=3936)
Records excluded (n=3562)
Records sought for retrieval (n=596)
Records not retrieved (n=18)
Records sought for retrieval (n=374)
Records not retrieved (n=20)
Full text records assessed for eligibility (n=578)
Records excluded (n=562):
Due to study design (n=220)
Due to intervention (n=172)
Due to abstract/commentary/ protocol (n=70)
Due to population (n=67)
Due to outcome measure (n=15)
Due to insufficient data (n=8)
Duplicate (n=6)
Due to no author reply (n=2)
Other (n=2)
Records assessed for eligibility (n= 354)
Records excluded: (n=341)
Due to study design (n=77)
Due to intervention (n=133)
Due to abstract/commentary/ protocol (n=12)
Due to population (n=41)
Due to outcome measure (n=5)
Due to insufficient data (n=14)
Duplicate (n=38)
Due to no author reply (n=6)
Other (n=2)
Duplicates of included database records (n=13)
Records from databases included (n=16)
Records from grey literature included (n= 13)
Total records included in review (n =29)
Reporting on 24 studies

Figure 2 rev

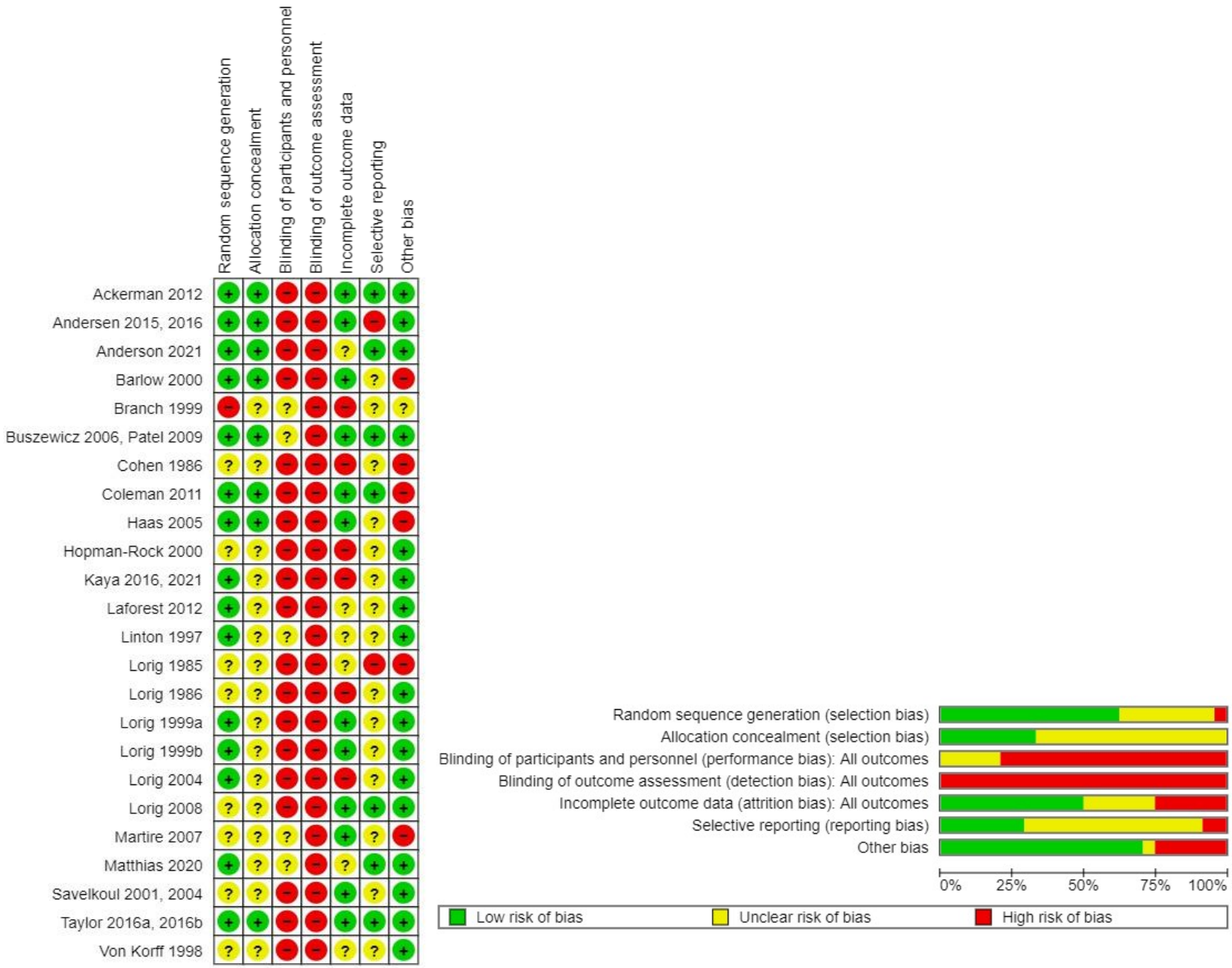
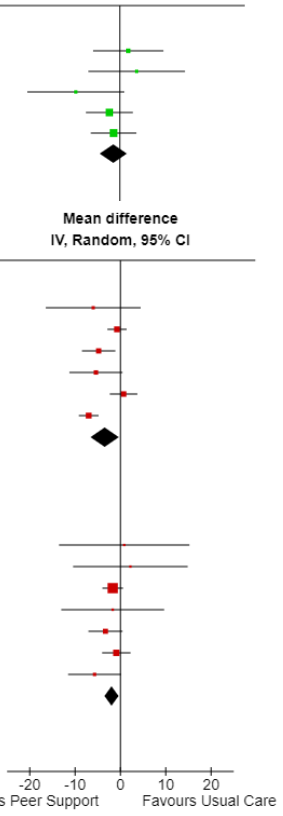

Random sequence generation
Allocation concealment
Blinding of participants and personnel
Blinding of outcome assessment
Incomplete outcome data
Selective reporting
Other bias
Ackerman 2012
Andersen 2015, 2016
Anderson 2021
Barlow 2000
Branch 1999
Buszewicz 2006, Patel 2009
Cohen 1986
Coleman 2011
Haas 2005
Hopman-Rock 2000
Kaya 2016, 2021
Laforest 2012
Linton 1997
Lorig 1985
Lorig 1986
Lorig 1999a
Lorig 1999b
Lorig 2004
Lorig 2008
Martire 2007
Matthias 2020
Savelkoul 2001, 2004
Taylor 2016a, 2016b
Von Korff 1998
Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias): All outcomes
Blinding of outcome assessment (detection bias): All outcomes
Incomplete outcome data (attrition bias): All outcomes
Selective reporting (reporting bias)
Other bias
0%
25%
50%
75%
100%
Low risk of bias
Unclear risk of bias
High risk of bias

Figure 3

a)

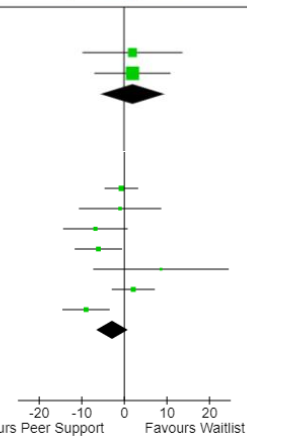
Study or Subgroup	Peer Support Mean	SD	Total	Usual Care Mean	SD	Total	Weight	Mean difference IV, Random, 95% CI
Short-Term								
Ackerman 2012*	38.2	22.2	40	36.2	18.1	55	15.0%	2.00 [-6.38 , 10.38]
Andersen 2015	60.6	26.3	46	56.6	29.9	46	7.9%	4.00 [-7.51 , 15.51]
Anderson 2021	37.5	19.5	22	48	20.5	24	7.9%	-10.50 [-22.06 , 1.06]
Martire 2007†	40.6	16.5	89	43.1	16.6	54	33.6%	-2.50 [-8.10 , 3.10]
Von Korff 1998	38.7	22.1	124	40.2	21.3	121	35.6%	-1.50 [-6.93 , 3.93]
Subtotal (95% CI)			**321**			**300**	**100.0%**	**-1.58 [-4.83 , 1.66]**
Heterogeneity: Tau² = 0.00; Chi² = 4.00, df = 4 (P = 0.41); I² = 0%								
Test for overall effect: Z = 0.96 (P = 0.34)								

Study or Subgroup	MD	SE	Peer Support Total	Usual Care Total	Weight	Mean difference IV, Random, 95% CI
Medium-Term						
Anderson 2021	-6	5.34	22	24	6.5%	-6.00 [-16.47 , 4.47]
Buszewicz 2006	-0.75	1.08	406	406	21.6%	-0.75 [-2.87 , 1.37]
Lorig 2008	-4.8	1.88	310	331	18.0%	-4.80 [-8.48 , -1.12]
Martire 2007†	-5.4	2.99	89	54	13.0%	-5.40 [-11.26 , 0.46]
Taylor 2016*	0.7	1.55	403	300	19.5%	0.70 [-2.34 , 3.74]
Von Korff 1998	-7	1.11	119	109	21.4%	-7.00 [-9.18 , -4.82]
Subtotal (95% CI)			**1349**	**1224**	**100.0%**	**-3.48 [-6.61 , -0.35]**
Heterogeneity: Tau² = 10.64; Chi² = 24.73, df = 5 (P = 0.0002); I² = 80%						
Test for overall effect: Z = 2.18 (P = 0.03)						
Long-Term						
Ackerman 2012*	0.84	7.34	38	56	1.2%	0.84 [-13.55 , 15.23]
Andersen 2016	2.2	6.45	46	46	1.5%	2.20 [-10.44 , 14.84]
Buszewicz 2006	-1.7	1.16	406	406	46.7%	-1.70 [-3.97 , 0.57]
Linton 1997	-1.7	5.79	39	25	1.9%	-1.70 [-13.05 , 9.65]
Lorig 2008	-3.3	1.92	307	344	17.1%	-3.30 [-7.06 , 0.46]
Taylor 2016*	-0.9	1.6	403	300	24.6%	-0.90 [-4.04 , 2.24]
Von Korff 1998	-5.7	2.98	112	106	7.1%	-5.70 [-11.54 , 0.14]
Subtotal (95% CI)			**1351**	**1283**	**100.0%**	**-1.97 [-3.53 , -0.42]**
Heterogeneity: Tau² = 0.00; Chi² = 3.11, df = 6 (P = 0.79); I² = 0%						
Test for overall effect: Z = 2.49 (P = 0.01)						



b)

Study or Subgroup	Peer Support Mean	SD	Total	Waitlist Control Mean	SD	Total	Weight	Mean difference IV, Random, 95% CI
Short-Term								
Cohen 1986	44.8	22.3	28	42.8	24.9	34	36.7%	2.00 [-9.76 , 13.76]
Hopman-Rock 2000*	27.2	21.4	55	25.2	23.5	44	63.3%	2.00 [-6.96 , 10.96]
Subtotal (95% CI)			**83**			**78**	**100.0%**	**2.00 [-5.12 , 9.12]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%								
Test for overall effect: Z = 0.55 (P = 0.58)								
Medium-Term‡								
Barlow 2000	-3.1	22.5	311	-2.4	23.8	233	20.9%	-0.70 [-4.65 , 3.25]
Haas 2005	-7.7	26	54	-6.7	23.6	47	9.6%	-1.00 [-10.68 , 8.68]
Hopman-Rock 2000*	1.7	19.9	55	8.5	18.5	44	12.8%	-6.80 [-14.39 , 0.79]
Lorig 1985	-10.4	21.1	129	-4.3	16.8	61	17.0%	-6.10 [-11.67 , -0.53]
Lorig 1986	-2.1	30.7	27	-10.7	30	29	4.5%	8.60 [-7.32 , 24.52]
Lorig 1999a	-5.4	15.9	86	-7.5	15.2	62	18.2%	2.10 [-2.96 , 7.16]
Lorig 1999b	-8.8	24	189	0.2	22	97	17.0%	-9.00 [-14.56 , -3.44]
Subtotal (95% CI)			**851**			**573**	**100.0%**	**-2.90 [-6.62 , 0.81]**
Heterogeneity: Tau² = 13.08; Chi² = 13.99, df = 6 (P = 0.03); I² = 57%								
Test for overall effect: Z = 1.53 (P = 0.13)								



c)

Study or Subgroup	Peer Support Mean	SD	Total	Active Control Mean	SD	Total	Weight	Mean difference IV, Random, 95% CI
Short-Term								
Andersen 2015	60.6	26.3	46	47.8	29.4	44	17.2%	12.80 [1.26 , 24.34]
Cohen 1986 §	44.8	22.3	28	40	21.7	24	16.1%	4.80 [-7.18 , 16.78]
Coleman 2011	33.8	15.5	83	30.8	16.2	85	66.6%	3.00 [-1.79 , 7.79]
Subtotal (95% CI)			**157**			**153**	**100.0%**	**4.98 [-0.08 , 10.04]**
Heterogeneity: Tau² = 4.04; Chi² = 2.36, df = 2 (P = 0.31); I² = 15%								
Test for overall effect: Z = 1.93 (P = 0.05)								
Medium-Term								
Coleman 2011	31.3	16.9	79	29.2	17.4	79	47.5%	2.10 [-3.25 , 7.45]
Lorig 1986	63.6	36.1	27	52.9	35.2	29	3.9%	10.70 [-8.00 , 29.40]
Matthias 2020§	58	19	92	57	17	86	48.6%	1.00 [-4.29 , 6.29]
Subtotal (95% CI)			**198**			**194**	**100.0%**	**1.90 [-1.79 , 5.59]**
Heterogeneity: Tau² = 0.00; Chi² = 0.97, df = 2 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.01 (P = 0.31)								
Long-Term								
Andersen 2016	53.8	29.9	46	41.5	29.9	44	5.5%	12.30 [-0.06 , 24.66]
Coleman 2011	30.9	15.8	77	26.3	16.5	79	28.4%	4.60 [-0.47 , 9.67]
Linton 1997	15.5	7.4	39	14	8.75	39	48.6%	1.50 [-2.10 , 5.10]
Lorig 2004§	39.3	28	142	38	32	166	17.5%	1.30 [-5.40 , 8.00]
Subtotal (95% CI)			**304**			**328**	**100.0%**	**2.94 [-0.01 , 5.90]**
Heterogeneity: Tau² = 1.30; Chi² = 3.45, df = 3 (P = 0.33); I² = 13%								
Test for overall effect: Z = 1.95 (P = 0.05)								

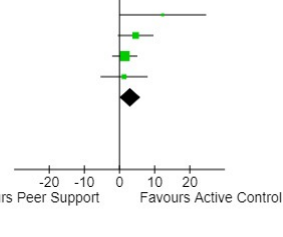


Figure 4

a)

Study or Subgroup	Peer Support Mean	SD	Total	Usual Care Mean	SD	Total	Weight	Std. mean difference IV, Random, 95% CI
Short-Term								
Anderson 2021	5.5	2.2	22	6.5	2.4	24	45.4%	-0.43 [-1.01 , 0.16]
Martire 2007*	148.8	28	89	138.3	28.1	54	54.6%	0.37 [0.03 , 0.71]
Subtotal (95% CI)			**111**			**78**	**100.0%**	**0.01 [-0.77 , 0.79]**
Heterogeneity: Tau² = 0.26; Chi² = 5.34, df = 1 (P = 0.02); I² = 81%								
Test for overall effect: Z = 0.03 (P = 0.98)								
Medium-Term								
Anderson 2021	6.3	1.6	22	5.3	2.5	24	3.0%	0.46 [-0.12 , 1.05]
Lorig 2008	5.7	2.1	310	5.1	2.1	331	42.4%	0.29 [0.13 , 0.44]
Martire 2007*	150.2	27.5	89	139.5	27.4	54	8.8%	0.39 [0.05 , 0.73]
Taylor 2016†	35.5	14	403	32.7	15	300	45.8%	0.19 [0.04 , 0.34]
Subtotal (95% CI)			**824**			**709**	**100.0%**	**0.26 [0.16 , 0.36]**
Heterogeneity: Tau² = 0.00; Chi² = 1.85, df = 3 (P = 0.60); I² = 0%								
Test for overall effect: Z = 4.98 (P < 0.00001)								
Long-Term								
Lorig 2008	5.9	2.1	307	5.3	2.1	344	49.1%	0.29 [0.13 , 0.44]
Taylor 2016†	35.4	14.1	403	33.4	15.1	300	50.9%	0.14 [-0.01 , 0.29]
Subtotal (95% CI)			**710**			**644**	**100.0%**	**0.21 [0.07 , 0.36]**
Heterogeneity: Tau² = 0.00; Chi² = 1.82, df = 1 (P = 0.18); I² = 45%								
Test for overall effect: Z = 2.84 (P = 0.005)								

Std. mean difference IV, Random, 95% CI

-1 -0.5 0 0.5 1

Favours Usual Care Favours Peer Support

b)

Study or Subgroup	Peer Support Mean	SD	Total	Waitlist Control Mean	SD	Total	Weight	Std. mean difference IV, Random, 95% CI
Medium-Term‡								
Barlow 2000	4.8	9.6	311	1.5	8.7	233	46.8%	0.36 [0.19 , 0.53]
Haas 2005	0.45	23	44	-0.5	24.8	39	11.5%	0.04 [-0.39 , 0.47]
Hopman-Rock 2000†	2	21.6	54	-3.5	17.4	44	13.1%	0.28 [-0.12 , 0.68]
Lorig 1999b	1.1	2.2	189	-0.04	2.2	97	28.6%	0.52 [0.27 , 0.77]
Subtotal (95% CI)			**598**			**413**	**100.0%**	**0.36 [0.20 , 0.51]**
Heterogeneity: Tau² = 0.01; Chi² = 3.83, df = 3 (P = 0.28); I² = 22%								
Test for overall effect: Z = 4.51 (P < 0.00001)								

Std. mean difference IV, Random, 95% CI

-1 -0.5 0 0.5 1

Favours Waitlist Favours Peer Support

c)

Study or Subgroup	Peer Support Mean	SD	Total	Active Control Mean	SD	Total	Weight	Mean difference IV, Random, 95% CI
Medium-Term								
Coleman 2011	6.6	1.7	79	6.9	1.7	79	57.1%	-0.30 [-0.83 , 0.23]
Matthias 2020§	6.3	2.3	92	6.1	2.1	86	42.9%	0.20 [-0.45 , 0.85]
Subtotal (95% CI)			**171**			**165**	**100.0%**	**-0.09 [-0.57 , 0.40]**
Heterogeneity: Tau² = 0.03; Chi² = 1.37, df = 1 (P = 0.24); I² = 27%								
Test for overall effect: Z = 0.35 (P = 0.73)								
Long-Term								
Coleman 2011	6.6	1.7	77	6.9	1.7	79	45.3%	-0.30 [-0.83 , 0.23]
Lorig 2004§	6.6	2.3	142	7.1	2	166	54.7%	-0.50 [-0.99 , -0.01]
Subtotal (95% CI)			**219**			**245**	**100.0%**	**-0.41 [-0.77 , -0.05]**
Heterogeneity: Tau² = 0.00; Chi² = 0.30, df = 1 (P = 0.59); I² = 0%								
Test for overall effect: Z = 2.23 (P = 0.03)								

Mean difference IV, Random, 95% CI

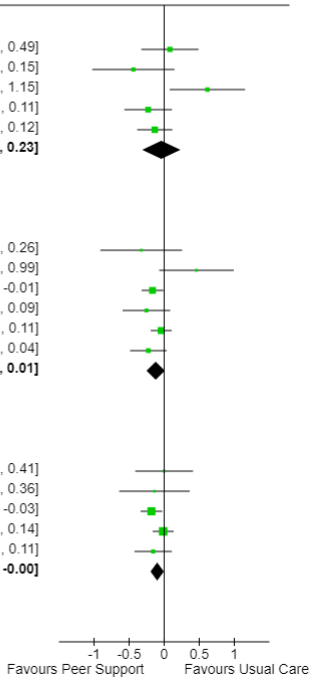
-1 -0.5 0 0.5 1

Favours Active Control Favours Peer Support

Figure 5

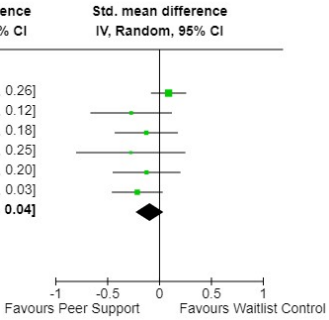
a)

Study or Subgroup	Peer Support Mean	SD	Total	Usual Care Mean	SD	Total	Weight	Std. mean difference IV, Random, 95% CI
Short-Term								
Ackerman 2012*	40.1	20.9	40	38.4	20.2	55	20.1%	0.08 [-0.33 , 0.49]
Anderson 2021	23.2	13.2	22	29.6	15.3	24	13.6%	-0.44 [-1.02 , 0.15]
Kaya 2016	3.7	2.1	27	2.51	1.7	29	15.1%	0.62 [0.08 , 1.15]
Martire 2007†	24.7	10.9	89	27.2	10.9	54	23.3%	-0.23 [-0.57 , 0.11]
Von Korff 1998	6.6	5.6	124	7.4	6.3	121	27.9%	-0.13 [-0.38 , 0.12]
Subtotal (95% CI)			**302**			**283**	**100.0%**	**-0.04 [-0.31 , 0.23]**
Heterogeneity: Tau² = 0.05; Chi² = 9.41, df = 4 (P = 0.05); I² = 57%								
Test for overall effect: Z = 0.29 (P = 0.77)								
22.2.2 Medium-Term								
Anderson 2021	23.6	14.8	22	28.3	13.5	24	4.6%	-0.33 [-0.91 , 0.26]
Kaya 2016	3.9	2.3	27	2.8	2.4	29	5.4%	0.46 [-0.07 , 0.99]
Lorig 2008	2	1.3	310	2.2	1.1	331	30.2%	-0.17 [-0.32 , -0.01]
Martire 2007†	23.1	11.4	89	26	11.4	54	11.6%	-0.25 [-0.59 , 0.09]
Taylor 2016*	53.2	25.7	403	54.3	26.7	300	31.1%	-0.04 [-0.19 , 0.11]
Von Korff 1998	5.8	5.9	119	7.2	6.5	109	17.0%	-0.23 [-0.49 , 0.04]
Subtotal (95% CI)			**970**			**847**	**100.0%**	**-0.12 [-0.25 , 0.01]**
Heterogeneity: Tau² = 0.01; Chi² = 7.67, df = 5 (P = 0.18); I² = 35%								
Test for overall effect: Z = 1.81 (P = 0.07)								
22.2.3 Long-Term								
Ackerman 2012*	40.7	20.2	38	40.7	13.1	56	5.3%	0.00 [-0.41 , 0.41]
Linton 1997	-3	0.69	39	-2.9	0.75	25	3.6%	-0.14 [-0.64 , 0.36]
Lorig 2008	1.9	1.2	307	2.1	1	344	38.0%	-0.18 [-0.34 , -0.03]
Taylor 2016*	52.9	28	403	53.3	28.8	300	40.4%	-0.01 [-0.16 , 0.14]
Von Korff 1998	5.8	6.3	112	6.8	6.4	106	12.8%	-0.16 [-0.42 , 0.11]
Subtotal (95% CI)			**899**			**831**	**100.0%**	**-0.10 [-0.19 , -0.00]**
Heterogeneity: Tau² = 0.00; Chi² = 2.77, df = 4 (P = 0.60); I² = 0%								
Test for overall effect: Z = 2.06 (P = 0.04)								



b)

Study or Subgroup	Peer Support Mean	SD	Total	Waitlist Control Mean	SD	Total	Weight	Std. mean difference IV, Random, 95% CI
Medium-Term‡								
Barlow 2000	0.01	0.34	311	-0.02	0.35	233	33.7%	0.09 [-0.08 , 0.26]
Haas 2005	-12.2	30.1	54	-4.2	27.7	47	10.1%	-0.27 [-0.67 , 0.12]
Lorig 1985	-0.06	0.34	129	-0.02	0.24	61	15.3%	-0.13 [-0.43 , 0.18]
Lorig 1986	-0.2	0.3	27	-0.1	0.4	29	6.0%	-0.28 [-0.80 , 0.25]
Lorig 1999 a	-0.05	0.38	86	0	0.42	62	13.7%	-0.13 [-0.45 , 0.20]
Lorig 1999 b	-0.1	0.49	189	0	0.41	97	21.2%	-0.21 [-0.46 , 0.03]
Subtotal (95% CI)			**796**			**529**	**100.0%**	**-0.10 [-0.23 , 0.04]**
Heterogeneity: Tau² = 0.01; Chi² = 6.49, df = 5 (P = 0.26); I² = 23%								
Test for overall effect: Z = 1.42 (P = 0.16)								



c)

Study or Subgroup	Peer Support Mean	SD	Total	Active Control Mean	SD	Total	Weight	Std. mean difference IV, Random, 95% CI
Short-Term								
Cohen 1986**	1.3	0.71	28	1.28	0.68	24	18.7%	0.03 [-0.52 , 0.57]
Coleman 2011	21.6	11.4	83	20.4	11.8	85	60.6%	0.10 [-0.20 , 0.41]
Laforest 2012§	3.3	0.74	26	3.3	0.68	32	20.7%	0.00 [-0.52 , 0.52]
Subtotal (95% CI)			**137**			**141**	**100.0%**	**0.07 [-0.17 , 0.30]**
Heterogeneity: Tau² = 0.00; Chi² = 0.14, df = 2 (P = 0.93); I² = 0%								
Test for overall effect: Z = 0.56 (P = 0.57)								
Medium-Term								
Coleman 2011	20.9	12.4	79	21.2	12.7	79	40.5%	-0.02 [-0.34 , 0.29]
Lorig 1986	0.8	0.8	27	0.7	0.7	29	14.3%	0.13 [-0.39 , 0.66]
Matthias 2020**	-47.3	24.4	92	-41.5	22.8	86	45.2%	-0.24 [-0.54 , 0.05]
Subtotal (95% CI)			**198**			**194**	**100.0%**	**-0.10 [-0.30 , 0.10]**
Heterogeneity: Tau² = 0.00; Chi² = 1.90, df = 2 (P = 0.39); I² = 0%								
Test for overall effect: Z = 1.00 (P = 0.32)								
Long-Term								
Coleman 2011	21.1	11.9	77	18.7	12.4	79	28.8%	0.20 [-0.12 , 0.51]
Laforest 2012§	3.2	0.56	21	3.6	0.77	29	14.3%	-0.57 [-1.14 , 0.00]
Linton 1997	-3	0.69	39	-3	0.94	39	20.1%	0.00 [-0.44 , 0.44]
Lorig 2004**	0.88	0.62	142	0.78	0.65	166	36.8%	0.16 [-0.07 , 0.38]
Subtotal (95% CI)			**279**			**313**	**100.0%**	**0.03 [-0.22 , 0.29]**
Heterogeneity: Tau² = 0.03; Chi² = 6.03, df = 3 (P = 0.11); I² = 50%								
Test for overall effect: Z = 0.25 (P = 0.80)								

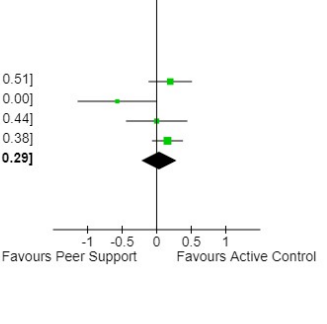
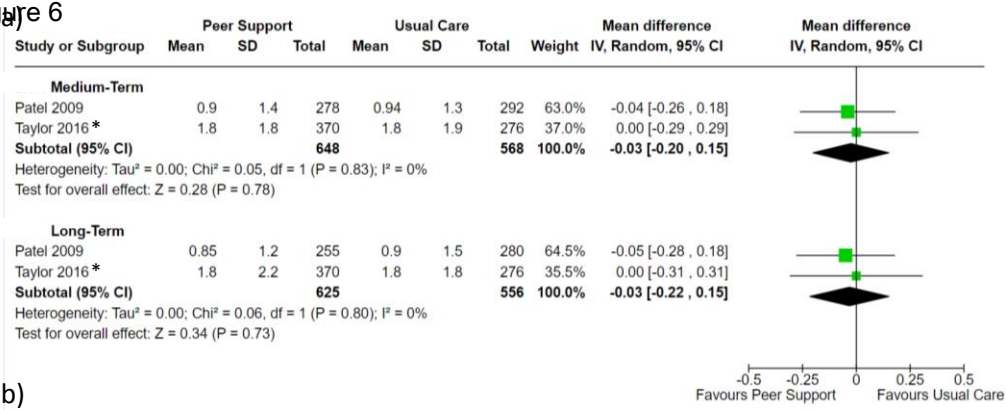
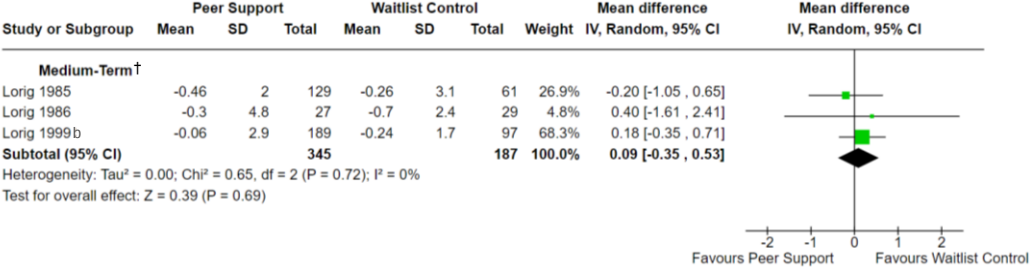


Figure 6

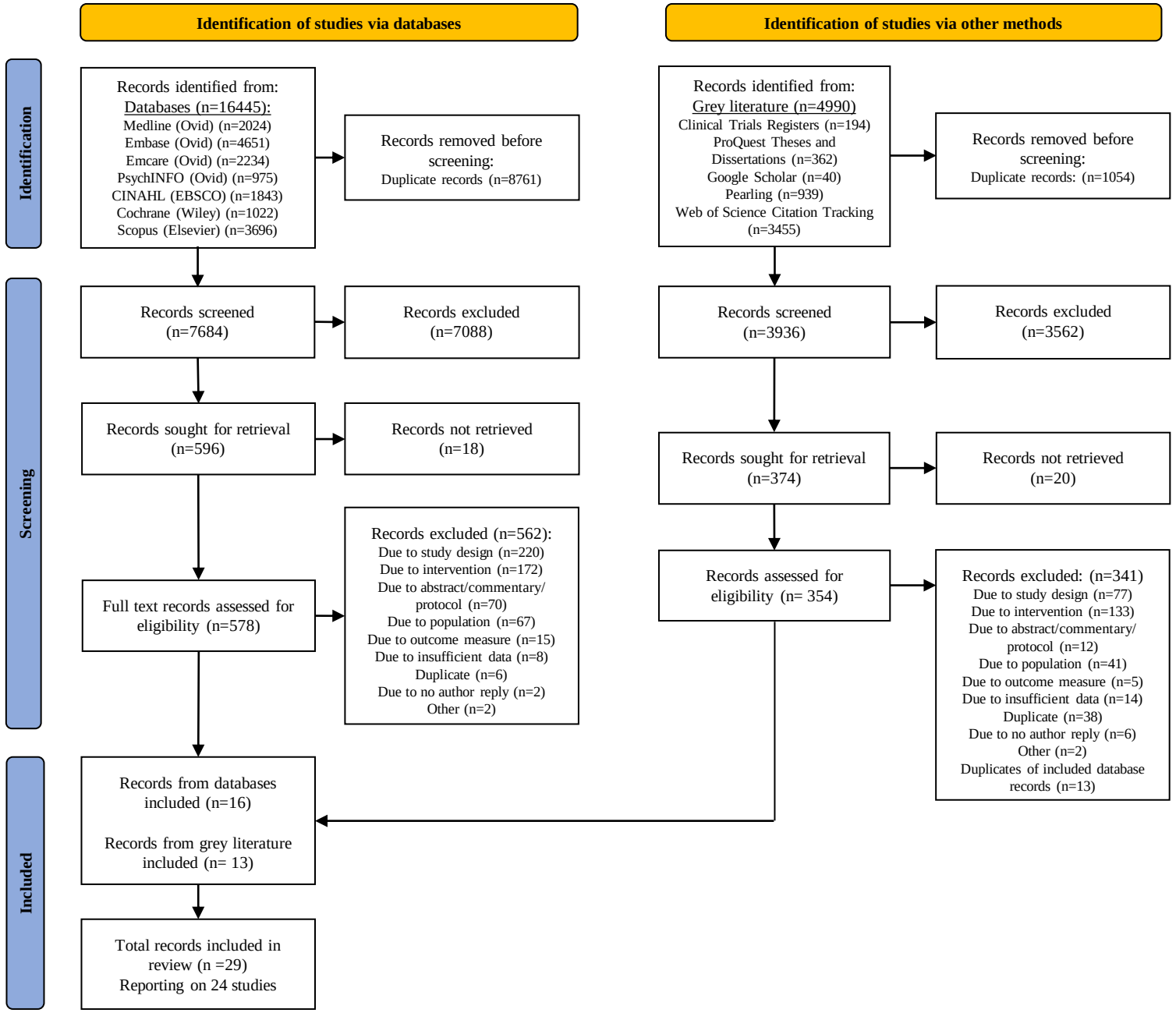
a)


Peer Support
Usual Care
Mean difference
Mean difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
Medium-Term
Patel 2009 0.9 1.4 278 0.94 1.3 292 63.0% -0.04 [-0.26 , 0.18]
Taylor 2016 * 1.8 1.8 370 1.8 1.9 276 37.0% 0.00 [-0.29 , 0.29]
Subtotal (95% CI) 648 568 100.0% -0.03 [-0.20 , 0.15]
Heterogeneity: Tau² = 0.00; Chi² = 0.05, df = 1 (P = 0.83); I² = 0%
Test for overall effect: Z = 0.28 (P = 0.78)
Long-Term
Patel 2009 0.85 1.2 255 0.9 1.5 280 64.5% -0.05 [-0.28 , 0.18]
Taylor 2016 * 1.8 2.2 370 1.8 1.8 276 35.5% 0.00 [-0.31 , 0.31]
Subtotal (95% CI) 625 556 100.0% -0.03 [-0.22 , 0.15]
Heterogeneity: Tau² = 0.00; Chi² = 0.06, df = 1 (P = 0.80); I² = 0%
Test for overall effect: Z = 0.34 (P = 0.73)
-0.5 -0.25 0 0.25 0.5
Favours Peer Support
Favours Usual Care

b)


Peer Support
Waitlist Control
Mean difference
Mean difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
Medium-Term†
Lorig 1985 -0.46 2 129 -0.26 3.1 61 26.9% -0.20 [-1.05 , 0.65]
Lorig 1986 -0.3 4.8 27 -0.7 2.4 29 4.8% 0.40 [-1.61 , 2.41]
Lorig 1999b -0.06 2.9 189 -0.24 1.7 97 68.3% 0.18 [-0.35 , 0.71]
Subtotal (95% CI) 345 187 100.0% 0.09 [-0.35 , 0.53]
Heterogeneity: Tau² = 0.00; Chi² = 0.65, df = 2 (P = 0.72); I² = 0%
Test for overall effect: Z = 0.39 (P = 0.69)
-2 -1 0 1 2
Favours Peer Support
Favours Waitlist Control

Consort/PRISMA Flowchart



PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	See Title Page
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	See Abstract
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Pg 1,2
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Pg 2
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Pg 3,4; Table 1
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Pg 3
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	See Appendix 1
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 2
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Pg 2,4
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Pg 4,5
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Pg 4
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Pg 4
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Pg 4,5
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	Pg 4,5
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	Pg 5
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Pg 5
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	Pg 4,5
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	Pg 5
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	Pg 5

PRISMA 2020 Checklist

Section and Topic	Item #	Checklist item	Location where item is reported
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	Pg 4,5
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	Pg 4
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Pg 6; Figure 1
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Figure 1
Study characteristics	17	Cite each included study and present its characteristics.	Pg 6; Table 2,3
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	Pg 7; Figure 2
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Pg 7-12; See Appendix 4,5
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	Pg 7-12
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Pg 7-12
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	Pg 7-12
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	Pg 12
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	Pg 7; Figure 2; See Appendix 3
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	Pg 7; See Appendix 3
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Pg 12-16
	23b	Discuss any limitations of the evidence included in the review.	Pg 15
	23c	Discuss any limitations of the review processes used.	Pg 15
	23d	Discuss implications of the results for practice, policy, and future research.	Pg 12-16
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	Pg 2
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	Pg 2
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	Pg 2; See Appendix 2
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Pg 17

Section and Topic	Item #	Checklist item	Location where item is reported
Competing interests	26	Declare any competing interests of review authors.	Pg 17
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Pg 3

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Appendix 1. Protocol Amendments.

The systematic review was prospectively registered on PROSPERO (CRD42022356850). Deviations from the protocol are listed below:

- Restricted systematic review from 'non-cancer chronic pain' to 'musculoskeletal chronic pain' population, to ensure narrow focus for body of evidence.
- Focussed analyses on the primary outcomes specifically delineated in our protocol (pain intensity, self-efficacy, function, health service utilisation, quality of life, perceived social support, and self-management), versus considering all reported outcomes; in addition to our primary outcomes. This was due to the numerous and varied outcomes evaluated in the included studies, with many studies reporting unique outcomes that could not be pooled.

Appendix 2. Search Strategies.

MEDLINE (via OVID)
exp Peer Group/
Mentors/
exp Mentoring/
Self-Help Groups/
((peer or peers) adj3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)).ti,ab,kf.
(peer2peer or peer to peer).ti,ab,kf.
(mentor* or pain coach or lay leader or lay led or train* volunteer or expert patient or volunteer leader).ti,ab,kf.
(consumer* adj2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)).ti,ab,kf.
(self help group* or selfhelp group* or support group* or therap* social club* or support commun*).ti,ab,kf.
((self manage* or selfmanage*) adj2 (program* or intervention or community based)).ti,ab,kf.
or/1-10
exp Pain/
(pain).ti,ab,kf.
or/12-13
11 and 14

Embase (via OVID)
exp peer group/
mentor/
exp mentoring/
exp support group/
((peer or peers) adj3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)).ti,ab,kf.
(peer2peer or peer to peer).ti,ab,kf.
(mentor* or pain coach or lay leader or lay led or train* volunteer or expert patient or volunteer leader).ti,ab,kf.
(consumer* adj2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)).ti,ab,kf.
(self help group* or selfhelp group* or support group* or therap* social club* or support commun*).ti,ab,kf.
((self manage* or selfmanage*) adj2 (program* or intervention or community based)).ti,ab,kf.
or/1-10
exp pain/
(pain).ti,ab,kf.
or/12-13
11 and 14

Emcare (via OVID)
exp peer group/
mentor/
exp mentoring/
exp support group/
((peer or peers) adj3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)).ti,ab,kf.
(peer2peer or peer to peer).ti,ab,kf.
(mentor* or pain coach or lay leader or lay led or train* volunteer or expert patient or volunteer leader).ti,ab,kf.
(consumer* adj2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)).ti,ab,kf.
(self help group* or selfhelp group* or support group* or therap* social club* or support commun*).ti,ab,kf.
((self manage* or selfmanage*) adj2 (program* or intervention or community based)).ti,ab,kf.
or/1-10
exp pain/
(pain).ti,ab,kf.
or/12-13
11 and 14

PsycInfo (via OVID)
Peers/
Mentor/
Support Groups/
((peer or peers) adj3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)).ti,ab,tw.
(peer2peer or peer to peer).ti,ab,tw.
(mentor* or pain coach or lay leader or lay led or train* volunteer or expert patient or volunteer leader).ti,ab,tw.
(consumer* adj2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)).ti,ab,tw.
(self help group* or selfhelp group* or support group* or therap* social club* or support commun*).ti,ab,tw.
((self manage* or selfmanage*) adj2 (program* or intervention or community based)).ti,ab,tw.
or/1-9
exp Pain/
(pain).ti,ab,tw.
or/11-12
10 and 13

CINAHL (via EBSCO)
MH "Peer Group"
MH "Mentorship"
MH "Peer Counseling"
MH "Support Groups"
((peer or peers) n3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)).TI OR AB.
(peer2peer or "peer to peer").TI OR AB.
(mentor* or "pain coach" or "lay leader" or "lay led" or "train* volunteer" or "expert patient" or "volunteer leader").TI OR AB.
(consumer* n2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)).TI OR AB.
("self help group*" or "selfhelp group*" or "support group*" or "therap* social club*" or "support commun*").TI OR AB.
(("self manage*" or selfmanage*) n2 (program* or intervention or "community based")).TI OR AB.
S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10
MH "Pain+"
(pain).TI OR AB.
S12 OR S13
S11 AND S14

The Cochrane Library (via Wiley)
[mh "Peer Group"]
[mh ^"Mentors"]
[mh "Mentoring"]
[mh "Self-Help Groups"]
((peer or peers) near/3 (support or specialist* or lead* or led or deliver* or direct* or provid* or conduct* or collaborat* or run or ran or organi* or manage* or operated or assist* or train* or companion or coach* or aide or volunteer* or work* or employ* or group or guide* or program or service* or influenc* or involv* or inclusion or participati*)):ti,ab,kw
(peer2peer or "peer to peer"):ti,ab,kw
(mentor* or "pain coach" or "lay leader" or "lay led" or train* NEXT volunteer or "expert patient" or "volunteer leader"):ti,ab,kw
(consumer* near/2 (support* or specialist* or service* or employ* or work* or provide* or traine* or run or ran or organi* or manage* or led or directed or delivered or operated or assist* or companion*)):ti,ab,kw
(self NEXT help NEXT group* or selfhelp NEXT group* or support NEXT group* or therap* NEXT social NEXT club* or support NEXT commun*):ti,ab,kw
((self NEXT manage* or selfmanage*) near/2 (program* or intervention or "community based")):ti,ab,kw
#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10
[mh "Pain"]
(pain):ti,ab,kw
#12 OR #13
#11 AND #14

Scopus (via Elsevier)

((TITLE-ABS-KEY ((peer OR peers) W/3 (support OR specialist* OR lead* OR led OR deliver* OR direct* OR provid* OR conduct* OR collaborat* OR run OR ran OR organi* OR manage* OR operated OR assist* OR train* OR companion OR coach* OR aide OR volunteer* OR work* OR employ* OR group OR guide* OR program OR service* OR influenc* OR involv* OR inclusion OR participati*))) OR (TITLE-ABS-KEY (peer2peer OR "peer to peer")) OR (TITLE-ABS-KEY (mentor* OR "pain coach" OR "lay leader" OR "lay led" OR "train* volunteer" OR "expert patient" OR "volunteer leader")) OR (TITLE-ABS-KEY (consumer* W/2 (support* OR specialist* OR service* OR employ* OR work* OR provide* OR traine* OR run OR ran OR organi* OR manage* OR led OR directed OR delivered OR operated OR assist* OR companion*))) OR (TITLE-ABS-KEY ("self help group*" OR "selfhelp group*" OR "support group*" OR "therap* social club*" OR "support commun*")) OR (TITLE-ABS-KEY (("self manage*" OR selfmanage*) W/2 (program* OR intervention OR "community based")))) AND (TITLE-ABS-KEY (pain))

Appendix 3. GRADE assessment of certainty in the evidence for each outcome.

Peer support vs. usual care Pain intensity: short-term										
		Certainty assessment					№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
5	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	321	300	MD: -1.58 [-4.83, 1.66], P=0.34, I^2=0%, N=5	◯⊕◯◯ Low
Peer support vs. usual care Pain intensity: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
6	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Very serious *We downgraded two levels due to high heterogeneity (I^2 value 80%).*	Not serious	Not serious	None	1309	1186	MD:-3.48 [-6.61, -0.35], P=0.03, I^2=80%, N=6	⊕◯◯◯ Very low
Peer support vs. usual care Pain intensity: long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
7	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	1322	1261	MD:-1.97 [-3.53, -0.42], P=0.01, I^2=0%, N=7	◯⊕◯◯ Low
Peer support vs. usual care Self-efficacy: short-term										

		Certainty assessment					№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Very serious *We downgraded two levels due to high heterogeneity (I^2 value 81%).*	Not serious	Very serious *Would downgrade one level as n<400 but already at very low.*	None	111	78	SMD 0.01 [-0.77, 0.79], P=0.98, I^2=81%, N=2	⊕◯◯◯ Very low

Peer support vs. usual care
Self-efficacy: medium-term

		Certainty assessment					№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
5	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	784	671	SMD:0.26 [0.16, 0.36], P<0.001, I^2=0%, N=4	◯⊕◯◯ Low

Peer support vs. usual care
Self-efficacy: long-term

		Certainty assessment					№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as one study was unable to determine the effect on peer support alone.*	Not serious	None	681	622	SMD:0.21 [0.07, 0.36], P=0.005, I^2=45%, N=2	⊕◯◯◯ Very low

Peer support vs. usual care
Function: short-term

		Certainty assessment					№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

5	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Serious *We downgraded one level due to moderate heterogeneity (I^2 value 57%).*	Not serious	Not serious	None	302	283	SMD: -0.04 [-0.31, 0.23], P=0.77, I^2=57%, N=5	⊕◯◯◯ Very low
Peer support vs. usual care Function: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
7	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	930	809	SMD:-0.12 [-0.25, 0.01], P=0.07, I^2=35%, N=6	◯⊕◯◯ Low
Peer support vs. usual care Function: long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
6	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as two studies were unable to determine the effect on peer support alone.*	Not serious	None	870	809	SMD:-0.10 [-0.19, -0.00], P=0.04, I^2=0%, N=5	⊕◯◯◯ Very low
Peer support vs. usual care Health service utilisation: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level due to differences in measures across health systems.*	Not serious	None	648	568	MD:-0.03 [-0.20, 0.15], P=0.78, I^2=0%, N=2	⊕◯◯◯ Very low
Peer support vs. usual care Health service utilisation: long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level due to differences in measures across health systems.*	Not serious	None	625	556	MD:-0.03 [-0.22, 0.15], P=0.73, I^2=0%, N=2	⊕◯◯◯ Very low
Peer support vs. usual care Quality of life: all timeframes										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as two studies were unable to determine the effect on peer support alone.*	Not serious	None	>400	>400	N/A	⊕◯◯◯ Very low
Peer support vs. usual care Social support: all timeframes										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

1	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	N/A *Not reported (single trial).*	Not serious	Serious *We downgraded one level as n<400.*	None	25	25	N/A	⊕◯◯◯ Very low

Peer support vs. usual care **Self-management: all timeframes**										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as one study was unable to determine the effect on peer support alone.*	Not serious	None	>400	>400	N/A	⊕◯◯◯ Very low

Peer support vs. waitlist control **Pain intensity: short-term**										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as one study was unable to determine the effect on peer support alone.*	Serious *Would downgrade one level (n<400) but already at very low.*	None	83	78	MD: 2.00 [-5.12, 9.12], P=0.58, I^2=0%, N=2	⊕◯◯◯ Very low

Peer support vs. waitlist control **Pain intensity: medium-term**										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

7	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Serious *We downgraded one level due to moderate heterogeneity (I^2 value 57%).*	Serious *Would downgrade one level as majority studies by same author (same program) but already at very low.*	Not serious	None	851	573	MD:-2.90 [-6.62, 0.81], P=0.13, I^2=57%, N=7	⊕◯◯◯ Very low
Peer support vs. waitlist control Self-efficacy: short-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
1	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	N/A *Not reported (single trial).*	Serious *We downgraded one level as study was unable to determine the effect on peer support alone.*	Serious *Would downgrade one level (n<400) but already at very low.*	None	60	60	N/A	⊕◯◯◯ Very low
Peer support vs. waitlist control Self-efficacy: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious *Only one study was unable to determine the effect on peer support alone, so not downgraded.*	Not serious	None	598	413	SMD:0.36 [0.20, 0.51], P<0.001, I^2=22%, N=4	◯⊕◯◯ Low
Peer support vs. waitlist control Function: short-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**

№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Peer support	Usual care	Absolute (95% CI)	
2	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	88	85	N/A	⊕◯◯◯ Very low
Peer support vs. waitlist control Function: medium-term										
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
7	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as majority studies by same author (same program).*	Not serious	None	796	529	SMD:-0.10 [-0.23, 0.04], P=0.16, I^2=23%, N=6	⊕◯◯◯ Very low
Peer support vs. waitlist control Health service utilisation: medium-term										
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
7	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as majority studies by same author (same program).*	Not serious	None	345	187	MD: 0.09 [-0.35, 0.53], P=0.69, I^2=0%, N=3	⊕◯◯◯ Very low
Peer support vs. waitlist control Quality of life: short- & medium-term										
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

1	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	N/A *Not reported (single trial).*	Serious *We downgraded one level as study was unable to determine the effect on peer support alone.*	Serious *Would downgrade one level (n<400) but already at very low.*	None	60	60	N/A	⊕◯◯◯ Very low

Peer support vs. waitlist control
Social support: all timeframes

			Certainty assessment				№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	<200	<200	N/A	⊕◯◯◯ Very low

Peer support vs. waitlist control
Self-management: all timeframes

			Certainty assessment				№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level as all studies only looked at one type of program.*	Not serious	None	>400	>400	N/A	⊕◯◯◯ Very low

Peer support vs. active control
Pain intensity: short-term

			Certainty assessment				№ of patients		Effect	Certainty
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	157	153	MD: 4.98 [-0.08, 10.04], P=0.05, I^2=15%, N=3	⊕◯◯◯ Very low
Peer support vs. active control Pain intensity: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	198	194	MD: 1.90 [-1.79, 5.59], P=0.31, I^2=0%, N=3	⊕◯◯◯ Very low
Peer support vs. active control Pain intensity: long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	304	328	MD: 2.94 [-0.01, 5.90], P=0.05, I^2=13%, N=4	◯⊕◯◯ Low
Peer support vs. active control Self-efficacy: short-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

1	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	N/A *Not reported (single trial).*	Not serious	Serious *We downgraded one level as n<400.*	None	90	90	N/A	⊕◯◯◯ Very low
Peer support vs. active control Self-efficacy: medium-term										
Certainty assessment							**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	171	165	MD: -0.09 [-0.57, 0.40], P=0.73, I^2=27%, N=2	⊕◯◯◯ Very low
Peer support vs. active control Self-efficacy: long-term										
Certainty assessment							**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	219	245	MD:-0.41 [-0.77, -0.05], P=0.03, I^2=0%, N=2	◯⊕◯◯ Low
Peer support vs. active control Function: short-term										
Certainty assessment							**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

3	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	137	141	SMD:0.07 [-0.17, 0.30], P=0.57, I^2=0%, N=3	⊕◯◯◯ Very low
Peer support vs. active control Function: medium-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	198	194	SMD -0.10 [-0.30, 0.10], P=0.32, I^2=0%, N=3	⊕◯◯◯ Very low
Peer support vs. active control Function: long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
4	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Not serious	None	279	313	SMD:0.03 [-0.22, 0.29], P=0.80, I^2=50%, N=4	◯⊕◯◯ Low
Peer support vs. active control Health service utilisation: medium- & long-term										
		Certainty assessment					**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

3	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Serious *We downgraded one level due to differences in measures across health systems.*	Not serious	None	>300	>300	N/A	⊕◯◯◯ Very low
				Peer support vs. active control Quality of life: medium- & long-term						
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trial	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	<200	<200	N/A	⊕◯◯◯ Very low
				Peer support vs. active control Social support: all timeframes						
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	
2	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	Not serious	Not serious	Serious *We downgraded one level as n<400.*	None	151	130	N/A	⊕◯◯◯ Very low
				Peer support vs. active control Self-management: all timeframes						
			Certainty assessment				**№ of patients**		**Effect**	**Certainty**
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Peer support**	**Usual care**	**Absolute (95% CI)**	

1	Randomised trials	Very serious *We downgraded two levels due to unclear or high risk of bias in multiple domains.*	N/A *Not reported (single trial)*	Not serious	Serious *We downgraded one level as n<400.*	None	120	95	N/A	⊕◯◯◯ Very low

Appendix 4. Summary of unpooled data for each outcome.

Control	Timepoint	Study	Effect size [95%CI], P-values if available
Pain intensity. *(Negative values favour peer support, 0-100 scale)*			
c. Active Control	Medium-Term	Matthias, 2020	9-month MD 1.00 [-4.31, 6.31]
	Long-Term	Lorig, 2004	2 year MD -3.33 [-10.97, 4.31], 3 year MD 4.00 [-8.36, 16.36]
Self-efficacy. *(Positive values favour peer support)*			
a. Usual Care	Short-Term	Branch, 1999	(Data no longer available)
	Medium-Term	Buszewicz, 2006	SMD 0.31 [-0.47, -0.14]*
	Long-Term	Buszewicz, 2006	Pain subscale adjusted difference in means 0.98 [0.07, 1.89]; Other subscale adjusted difference in means 1.58 [0.25, 2.90]
b. Waitlist	Short-Term	Hopman-Rock 2000	SMD 0.40 [0.00, 0.81]
c. Active Control	Short-Term	Coleman, 2011	SMD -0.27 [-0.58, 0.03]
	Medium-Term	Matthias, 2020	9-month SMD 0.14 [-0.17, 0.44]
	Long-Term	Lorig, 2004	2 year SMD -0.18 [-0.42, 0.05], 3 year SMD -0.03 [-0.28, 0.22]
Function. *(Negative values favour peer support)*			
a. Usual Care	Medium-Term	Buszewicz, 2006	SMD -0.05 [-0.21, 0.11]*
	Long-Term	Buszewicz, 2006	SMD -0.07 [-0.23, 0.10]*
b. Waitlist	Short-Term	Cohen, 1986	(Data not available)
c. Active Control	Medium-Term	Matthias 2020	9-month SMD -0.14 [-0.44, 0.16]
	Long-Term	Lorig 2004	2 year SMD -0.06 [-0.30, 0.17], 3 year SMD -0.05 [-0.30, 0.20]
Health service utilisation. *(Negative values favour peer support)*			
a. Usual Care	Short-Term	Ackerman, 2012	4 measures: p-values range 0.21 to 0.92
	Medium-Term	Lorig, 2008	Physician visits SMD -0.14 [-0.29, 0.02]; Emergency visits SMD -0.04 [-0.20, 0.11]
		Anderson, 2021	14 measures (significant ones reported): Inpatient stays MD -0.91 [-1.78, -0.04]; GP practice nurse visits MD 0.93 [0.15, 1.71]
	Long-Term	Lorig, 2008	Physician visits SMD -0.09 [-0.24, 0.07]; Emergency visits SMD -0.02 [-0.17, 0.14]
		Ackerman, 2012	4 measures: p-values range 0.02 to 0.41
b. Waitlist	Medium-Term	Barlow, 2000	GP visits (other) SMD 0.16 [-0.14, 0.47]*; GP visits (arthritis) SMD -0.04 [-0.35, 0.27]*
		Haas, 2005	7 measures: p-values range 0.114 to 0.698
		Hopman-Rock, 2000	GP visits & medication effect size <0.2, p-value >0.05; Physical therapy visits effect size >0.4 *(favours peer support)*, p-value = 0.01
		Lorig, 1999a	Physician & ER visits SMD 0.21 [-0.11, 0.54]*; Hospital stays SMD -0.02 [-0.35, 0.31]*; Hospital nights SMD -0.26 [-0.59, 0.07]*
c. Active Control	Medium-Term	Lorig, 1986	Physician visits SMD 0.11 [-0.42, 0.63]*
		Matthias, 2020	Outpatient visits 9-month SMD 0.05 [-0.25, 0.36]; Emergency visits 9-month SMD -0.20 [-0.51, 0.10]
	Long-Term	Lorig, 2004	Physician visits 1 year SMD -0.24 [-0.47, -0.02]*, 2 year SMD 0.08 [-0.15, 0.32]*, 3 year SMD -0.33 [-0.59, -0.08]*; Rheumatologist visits 1 year SMD -0.08 [-0.30, 0.15]*, 2 year SMD 0.23 [-0.01, 0.46]*, 3 year SMD 0.07 [-0.18, 0.32]*

Quality of life. *(Positive values favour peer support)*			
a. Usual Care	Short-Term	Ackerman, 2011	SMD -0.28 [-0.69, 0.13]*
		Kaya, 2016	(No overall score data available)
	Medium-Term	Buszewicz, 2006/Patel, 2009	(No overall score data available)
		Kaya, 2016	(No overall score data available)
		Taylor, 2016	SMD 0.13 [−0.03, 0.29]
	Long-Term	Ackerman, 2012	SMD -0.11 [-0.52, 0.29]*
		Buszewicz, 2006/Patel, 2009	(No overall score data available)
		Taylor, 2016	SMD 0.01 [−0.16, 0.17]
b. Waitlist	Short-Term	Hopman-Rock, 2000	p-value < 0.05
	Medium-Term	Hopman-Rock, 2000	p-value > 0.05
c. Active Control	Short-Term	Coleman, 2011	(No overall score data available)
	Medium-Term	Coleman, 2011	(No overall score data available)
		Matthias, 2020	(No overall score data available)
	Long-Term	Coleman, 2011	(No overall score data available)
Social support. *(Positive values favour peer support)*			
a. Usual Care	Short-Term	Anderson, 2021	Effect size 0.0 [−0.80, 0.80]
	Medium-Term	Anderson, 2021	Effect size −0.40 [−0.84, 0.03]
b. Waitlist	Short-Term	Cohen, 1986	(Data not available)
	Medium-Term	Savelkoul, 2001/Savelkoul, 2004	(Data not available)
c. Active Control	Short-Term	Cohen, 1986	(Data not available)
	Medium-Term	Matthias, 2020	6-month SMD 0.09 [-0.20, 0.39], 9-month SMD 0.09 [-0.22, 0.39]
		Savelkoul, 2001/Savelkoul, 2004	(Data not available)
Self-management. *(Positive values favour peer support)*			
a. Usual Care	Short-Term	Ackerman, 2012	8 measures (significant reported here): Skill & technique acquisition adjusted between group difference 0.29 [0.04, 0.55] (high =better)
		Anderson, 2021	Effect size 8.3 [2.2, 14.4]
	Medium-Term	Anderson, 2021	Effect size 4.4 [−2.8, 11.6]
		Lorig, 2008	4 measures (all non-significant)
	Long-Term	Ackerman, 2012	8 measures (all non-significant)
		Lorig, 2008	4 measures (all non-significant)
b. Waitlist	Short-Term	Cohen, 1986	5 measures (all non-significant)
	Medium-Term	Barlow, 2000	Cognitive symptom management SMD 0.22 [0.05, 0.39] (p<0.0005)*; Relaxation (p<0.0005); flexibility exercises (p<0.0005); Strengthening exercises (p<0.0005)
		Lorig, 1985	Knowledge SMD 0.69 [0.38, 1.01]*; Arthritis exercise SMD 0.72 [0.41, 1.04]*; Relaxation SMD 0.64 [0.33, 0.95]*
		Lorig, 1999b	Range of motion exercise effect size 0.41 (p=0.004); Aerobic exercise effect size 0.07 (p=0.41)
c. Active Control	Short-Term	Cohen, 1986	5 measures (all non-significant)
	Medium-Term	Matthias, 2020	6-month SMD 0.33 [0.03, 0.62], 9-month SMD 0.15 [-0.15, 0.46]

*Calculated with within-group change data.

Appendix 5. Self-reported study outcomes summary.

Author, year	Intervention Group & Overview	Outcome measures* & follow-up timepoints
Studies with usual care controls only		
Ackerman, 2012	Int: Arthritis Self-Mx Program Con: usual care	Pain ◐, physical function ◐, quality of life ◐, self-management skills ▣, health service utilisation ◐, disease severity ◐, psychological distress ◐, stiffness ◐ *Baseline, 6 weeks, 3 months, 12 months*
Anderson, 2021	Int: one-to-one peer mentorship sessions Con: usual care	Pain ◐, function ◐, self-efficacy ◐, self-management ●, perceived social support ◐, healthcare & community resource use ◐, health status ◐, anxiety & depression ◐, stiffness ◐ *Baseline, 8 weeks, 6 months*
Branch, 1999	Int: arthritis patient educator Con: usual care	Arthritis impact (health status) ◐, self-efficacy ◐, arthritis knowledge ▣ *Baseline, 8 weeks*
Buszewicz, 2006 Patel, 2009	Int: arthritis self-Mx program Con: usual care	Pain ◐, quality of life ◐, physical functioning ◐, self-efficacy ●, anxiety & depression ● *Baseline, 4 months, 12 months* Quality of life ◐, resource use ▣, mental health ◐, physical health ◐, health state ◐ *Baseline, 4 months, 12 months*
Kaya, 2016 Kaya, 2021	Int: peer education group Con: usual care	Functional status ◐, activity status ◐, health status (multiple domains) ◐, quality of life ◐, depression ◐ *Baseline, 4 weeks, 6 months* Knowledge ◐ *Baseline, 4 weeks, 6 months*
Lorig, 2008	Int: internet-based arthritis self-Mx program Con: usual care	Pain ●, disability ◐, self-efficacy ●, health-related behaviors ◐, healthcare utilisation ◐, global health ●, fatigue ▣, health distress ●, activity limitations ● *Baseline, 6 months, 12 months*
Martire, 2007	Int A: patient education & support Int B: couples-orientated education & support Con: usual care	Pain ◐, physical function ◐, self-efficacy ◐, marital satisfaction ◐, stiffness ◐, depressive symptoms ◐ *Baseline, 6 weeks, 6 months*
Taylor, 2016a Taylor, 2016b	Int: self-Mx group course Con: usual care	Pain ◐, pain-related disability ◐, self-efficacy ●, social integration ●, health utility (quality of life) ◐, depression & anxiety ●, pain acceptance & coping ●, global health ◐ *Baseline, 6 months, 12 months* As above + healthcare utilisation ◐ *Baseline, 6 months, 12 months*
Von Korff, 1998	Int: lay-led self-Mx group Con: usual care	Pain ◐, pain interference ◐, impairment & limitation ●, self-care attitudes ●, back pain worries ●, mental health ◐

		Baseline, 3 months, 6 months, 12 months
Studies with waitlist controls only		
Barlow, 2000	Int: arthritis self-Mx program Con: wait-list control	Pain ◐, physical functioning ◐, self-efficacy ●, use of cognitive symptom management ●, health behaviors ▣, visits to GP ◐, health status ◐, fatigue ●, psychological wellbeing ● *Baseline, 4 months (12 months for 1 group only)*
Haas, 2005	Int: chronic disease self-Mx program Con: wait-list control	Pain ◐, functional disability ◐, self-efficacy ◐, health care utilisation ◐, general health ◐, emotional well-being ●, self-care attitudes ◐, pain days ◐, disability days ◐ *Baseline, 6 months*
Hopman-Rock, 2000	Int: living with OA program Con: wait-list control	Pain ▣, mobility ◐, quality of life ◐, self-efficacy ●, knowledge ●, health care utilisation ▣ *Baseline, 6 weeks, 6 months*
Lorig, 1985	Int: arthritis self-Mx program Con: wait-list control	Pain ●, physical disability ◐, self-management activities ●, health service utilisation ◐, knowledge ●, locus of control ◐ *Baseline, 4 months*
Lorig, 1999*	Int: chronic disease self-Mx program Con: wait-list control	Pain/physical discomfort ◐, disability ●, health service utilisation ▣, self-rated health ●, psychological wellbeing ◐, energy/fatigue ●, health distress ●, health behaviours ●, social/role activity limitations ●, shortness of breath ◐ *Baseline, 6 months*
Lorig, 1999 (SP)	Int: Spanish arthritis self-Mx program Con: wait-list control	Pain ●, disability ●, self-efficacy ●, self-management behaviour ▣, visits to physician ◐, self-rated health ◐, depression ◐, medication use ◐ *Baseline, 4 months*
Studies with active controls only		
Coleman, 2011	Int: lay-led arthritis self-Mx program Con: active control	Pain ○, physical function ○, quality of life ○, self-efficacy ◐, global health ◐ *Baseline, 8 weeks, 6 months, 12 months*
Lorig, 2004	Int: arthritis self-Mx program Con: active control	Pain ◐, disability ○, self-efficacy ○, healthcare utilisation ▣, depression ◐, global severity of arthritis ◐, role function ● *Baseline, 12 months, 24 months, 36 months*
Matthias, 2020	Int: one-to-one peer support Con: active control	Pain ◐, physical functioning ◐, quality of life ◐, self-efficacy ◐, self-management ◐, perceived social support ◐, healthcare utilisation ◐, general health perceptions ◐, anxiety & depression ◐, pain coping ◐ *Baseline, 6 months, 9 months*
Studies with multiple control groups		
Andersen, 2015	Int: chronic disease self-Mx program Con: usual care Other: active control	Pain ◐, work ability ◐, kinesiophobia ◐ *Baseline, 3 months*

Anderson, 2016		Pain ◐, work ability ◐, kinesiophobia ◐ *Baseline, 11 months*
Cohen, 1986	Int: lay-led arthritis self-Mx course Con: wait-list control Other: active control	Pain ◐, functional disability ◐, knowledge ◐, self-management behaviors ◐, perceived affective & instrumental support ◐, depression ◐ *Baseline, 6 weeks, 3 months*
Laforest, 2012	Int: social reinforcement with self-Mx intervention Con: active control Other: wait-list control	Functional limitations ●, coping ◐, helplessness ◐ *Baseline, 2 months, 10 months*
Linton, 1997	Int: lay-led educational support group Con: usual care Other: active control	Pain ◐, function ◐, coping strategies ◐, health status ◐, sick leave ◐, pain beliefs & attitudes ◐, overall outcome evaluation ◐ *Baseline, 12 months*
Lorig, 1986	Int: lay-led arthritis self-Mx course Con: wait-list control Other: active control	Pain ◐, disability ◐, self-management knowledge ○, self-management behaviours ●, visits to physician ◐, exercise frequency ▣ *Baseline, 4 months*
Savelkoul, 2001 Savelkoul, 2004†	Int: mutual support group Con: wait-list control Other: active control	Social interactions ◐, functional health status ◐, mobility ◐, action directed coping ○, loneliness ○, life satisfaction ○, coping by seeking social support ◐ *Baseline, 6 months* Social network size ◐, social skills ●, functional health status ◐, loneliness ◐, life satisfaction ◐ *Baseline, 6 months*

◐ = no between-group differences. ● = in favour of peer support intervention (statistically significant or assumed). ○ = in favour of control (statistically significant or assumed). ▣ = multiple domains with varied results, with at least one significant in favour of peer support.

*Outcomes reported for entire sample with all clinical conditions, not arthritis only. †Paper reported on mutual support group versus coping intervention group only. Int = intervention group; Cont = control group; Mx = management.

Appendix 6. Sensitivity analyses.

Table 1: Sensitivity analysis using correlation coefficient variation.

Use of correlation coefficients					
Paper	**Control**	**Outcome**	**Timepoint**	**Correlation coefficient (r)**	**MD or SMD [95% CI], I^2 value, (p-value)**
Ackerman 2012	Usual care	Pain	Short-term	0.67 (used)	-1.58 [4.83, 1.66], I^2 = 0%, (P = 0.34)
				0.77	-1.76 [-5.09, 1.56], I^2 = 0%, (P = 0.30)
				0.57	N/A
			Long-term	0.67 (used)	-1.97 [-3.53, -0.42], I^2 = 0%, (P = 0.01)
				0.77	-2.00 [-3.57, -0.44], I^2 = 0%, (P = 0.01)
				0.57	-2.00 [-3.57, -0.44], I^2 = 0%, (P = 0.02)
		Function	Short-term	0.6 (used)	-0.04 [-0.31, 0.23], I^2 = 57%, (P = 0.77)
				0.7	-0.04 [-0.31, 0.24], I^2 = 58%, (P = 0.78)
				0.5	N/A
			Long-term	0.6 (used)	-0.10 [-0.19, -0.00], I^2 = 0%, (P = 0.04)
				0.7	-0.10 [-0.19, -0.00], I^2 = 0%, (P = 0.04)
				0.5	-0.10 [-0.19, -0.00], I^2 = 0%, (P = 0.04)
Hopman-Rock 2000	Waitlist	Pain	Medium-term	0.58 (used)	-2.90 [-6.62, 0.81], I^2 = 57%, (P = 0.13)
				0.68	-2.98 [-6.66, 0.71], I^2 = 58%, (P = 0.11)
				0.48	-2.85 [-6.58, 0.89], I^2 = 56%, (P = 0.13)
		Self-efficacy	Medium-term	0.38 (used)	0.36 [0.20, 0.51], I^2 = 22%, (P < 0.001)
				0.48	0.36 [0.21, 0.51], I^2 = 20%, (P < 0.001)
				0.28	0.35 [0.19, 0.51], I^2 = 23%, (P < 0.001)
Lorig 1986	Active	Pain	Medium-term	0.6 (used)	1.90 [-1.79, 5.59], I^2 = 0%, (P = 0.31)
				0.7	1.80 [-1.91, 5.51], I^2 = 0%, (P = 0.34)
				0.5	2.08 [-1.57, 5.73], I^2 = 0%, (P = 0.26)
		Function	Medium-term	0.67 (used)*	-0.10 [-0.30, 0.10], I^2 = 0%, (P = 0.32)
				0.77	N/A
				0.57	N/A
Lorig 2004	Active	Pain	Long-term	0.6 (used)	2.94 [-0.01, 5.90], I^2 = 13%, (P = 0.05)
				0.7	3.01 [0.01, 6.02], I^2 = 12%, (P = 0.05)
				0.5	2.84 [-0.03, 5.72], I^2 = 15%, (P = 0.05)
		Self-efficacy	Long-term	0.6 (used)	-0.41 [-0.77, -0.05], I^2 = 0%, (P = 0.03)
				0.7	**-0.39 [-0.78, -0.00], I^2 = 0%, (P = 0.05)**
				0.5	-0.43 [-0.75, -0.11], I^2 = 0%, (P = 0.008)
		Function	Long-term	0.67 (used)*	0.03 [-0.22, 0.29], I^2 = 50%, (P = 0.80)
				0.77	0.03 [-0.22, 0.28], I^2 = 50%, (P = 0.81)
				0.57	N/A

*Calculation with correlation coefficient did not result in a real number (so imputed baseline SDs).

Table 2: Sensitivity analysis for standard deviation imputation.

Imputation of standard deviations					
Paper	**Control**	**Outcome**	**Timepoint**	**SD used**	**MD or SMD [95% CI], I^2 value, (p-value)**
Cohen 1986	Active	Pain	Short-term	(used)	4.98 [-0.08, 10.04], I^2=15%, (P=0.05)
				Removal of study	6.44 [-2.73, 15.60], I^2=58%, (P=0.17)
		Function	Short-term	(used)	0.07 [-0.17, 0.30], I^2=0%, (P=0.57)
				Removal of study	0.08 [-0.18, 0.34], I^2=0%, (P=0.56)
				Lowest SD	0.07 [-0.17, 0.30], I^2=0%, (P=0.57)
				Highest SD	0.07 [-0.17, 0.30], I^2=0%, (P=0.58)

Table 3: Sensitivity analysis for multi-component studies.

Effect on peer support alone cannot be seen					
Paper	**Control**	**Outcome**	**Timepoint**		**MD or SMD [95% CI], I^2 value, (p-value)**
Ackerman 2012	Usual care	Pain	Short-term	(used)	-1.58 [-4.83, 1.66], I^2=0%, (P=0.34)
				Removal of study	-2.23 [-5.90, 1.44], I^2=5%, (P=0.23)
		Function	Short-term	(used)	-0.04 [-0.31, 0.23], I^2=57%, (P=0.77)
				Removal of study	-0.07 [-0.41, 0.27], I^2=66%, (P=0.70)
Hopman-Rock 2000	Waitlist	Pain	Medium-term	(used)	-2.90 [-6.62, 0.81], I^2=57%, (P=0.13)
				Removal of study	-2.31 [-6.42, 1.80], I^2=61%, (P=0.27)
		Self-efficacy	Medium-term	(used)	0.36 [0.20, 0.51], I^2=22%, (P<0.001)
				Removal of study	0.36 [0.16, 0.56], I^2=45%, (P<0.005)
Taylor 2016	Usual care	Pain	Medium-term	(used)	-3.48 [-6.61, -0.35], I^2=80%, (P=0.03)
				Removal of study	-4.48 [-7.79, -1.16], I^2=76%, (P=0.008)
		Function	Medium-term	(used)	-0.12 [-0.25, 0.01], I^2=35%, (P=0.07)
				Removal of study	-0.15 [-0.32, 0.02], I^2=34%, (P=0.08)
		Self-efficacy	Medium-term	(used)	0.26 [0.16, 0.36], I^2=0%, (P<0.001)
				Removal of study	0.31 [0.17, 0.45], I^2=0%, (P<0.001)
Ackerman 2012 & Taylor 2016	Usual care	Pain	Long-term	(used)	-1.97 [-3.53, -0.42], I^2=0%, (P=0.01)
				Removal 2x studies	-2.37 [-4.17, -0.57], I^2=0%, (P=0.01)
		Function	Long-term	(used)	-0.10 [-0.19, -0.00], I^2=0%, (P=0.04)
				Removal 2x studies	-0.17 [-0.30, -0.04], I^2=0%, (P=0.009)

Table 4: Sensitivity analysis for imputing alternate peer group data.

Imputing alternate intervention group data					
Paper	**Control**	**Outcome**	**Timepoint**	**Data set used**	**SMD [95% CI], I^2 value, (p-value)**
Martire 2007	Usual care	Pain	Short-term	PES (used)	-1.58 [-4.83, 1.66], I^2=0%, (P=0.34)
				CES	-2.64 [-6.79, 1.50], I^2=33%, (P=0.21)
			Medium-term	PES (used)	-3.48 [-6.61, -0.35], I^2=80%, (P=0.03)
				CES	-3.31 [-6.40, -0.23], I^2=79%, (P=0.04)
		Self-efficacy	Short-term	PES (used)	0.01 [-0.77, 0.79], I^2=81%, (P=0.98)
				CES	-0.01 [-0.75, 0.73], I^2=79%, (P=0.98)
			Medium-term	PES (used)	0.26 [0.16, 0.36], I^2=0%, (P<0.001)
				CES	0.26 [0.16, 0.36], I^2=0%, (P<0.001)
		Function	Short-term	PES (used)	-0.04 [-0.31, 0.23], I^2=57%, (P=0.77)
				CES	-0.06 [-0.35, 0.23], I^2=62%, (P=0.69)
			Medium-term	PES (used)	-0.12 [-0.25, 0.01], I^2=35%, (P=0.07)
				CES	-0.11 [-0.24, 0.01], I^2=31%, (P=0.07)

Table 5: Sensitivity analysis for imputing alternate measure data for outcome.

Imputing alternate measure for outcome					
Paper	**Control**	**Outcome**	**Timepoint**	**Outcome measure used**	**SMD [95% CI], I^2 value, (p-value)**
Coleman 2011	Active	Pain	Short-term	WOMAC (used)	4.98 [-0.08, 10.04], I^2=15%, (P=0.05)
				SF-36 Bodily Pain	4.68 [-2.51, 11.87], I^2=44%, (P=0.20)
			Medium-term	WOMAC (used)	1.90 [-1.79, 5.59], I^2=0%, (P=0.31)
				SF-36 Bodily Pain	1.35 [-2.59, 5.29], I^2=0%, (P=0.50)
			Long-term	WOMAC (used)	2.94 [-0.01, 5.90], I^2=13%, (P=0.05)
				SF-36 Bodily Pain	1.85 [-0.90, 4.61], I^2=0%, (P=0.19)
		Function	Short-term	WOMAC (used)	0.07 [-0.17, 0.30], I^2=0%, (P=0.57)
				SF-36 Physical Function	0.08 [-0.16, 0.31], I^2=0%, (P=0.53)
			Medium-term	WOMAC (used)	-0.10 [-0.30, 0.10], I^2=0%, (P=0.32)
				SF-36 Physical Function	-0.02 [-0.29, 0.25], I^2=41%, (P=0.86)
			Long-term	WOMAC (used)	0.03 [-0.22, 0.29], I^2=50%, (P=0.80)
				SF-36 Physical Function	0.01 [-0.23, 0.25], I^2=45%, (P=0.92)
Lorig 1985	Waitlist	Pain	Medium-term	VAS (used)	-2.90 [-6.62, 0.81], I^2=57%, (P=0.13)
				Ordinal scale	-3.79 [-8.54, 0.97], I^2=71%, (P=0.12)