Systematic Review to Inform a World Health Organization (WHO) Clinical Practice Guideline: Benefits and Harms of Structured Exercise Programs for Chronic Primary Low Back Pain in Adults

Leslie Verville[1] · Rachel Ogilvie[2] · Cesar A. Hincapié[3,4,5] · Danielle Southerst[1] · Hainan Yu[1] · André Bussières[6,7] · Douglas P. Gross[8] · Paulo Pereira[9] · Silvano Mior[1,10] · Andrea C. Tricco[11,12,13] · Christine Cedraschi[14,15] · Ginny Brunton[1,16,17] · Margareta Nordin[18] · Gaelan Connell[1] · Jessica J. Wong[1] · Heather M. Shearer[1,10,19] · Joyce G. B. Lee[10] · Dan Wang[1] · Jill A. Hayden[2] · Carol Cancelliere[1]

Abstract

Purpose Evaluate benefits and harms of structured exercise programs for chronic primary low back pain (CPLBP) in adults to inform a World Health Organization (WHO) standard clinical guideline.

Methods We searched for randomized controlled trials (RCTs) in electronic databases (inception to 17 May 2022). Eligible RCTs targeted structured exercise programs compared to placebo/sham, usual care, or no intervention (including comparison interventions where the attributable effect of exercise could be isolated). We extracted outcomes, appraised risk of bias, conducted meta-analyses where appropriate, and assessed certainty of evidence using GRADE.

Results We screened 2503 records (after initial screening through Cochrane RCT Classifier and Cochrane Crowd) and 398 full text RCTs. Thirteen RCTs rated with overall low or unclear risk of bias were synthesized. Assessing individual exercise types (predominantly very low certainty evidence), pain reduction was associated with aerobic exercise and Pilates vs. no intervention, and motor control exercise vs. sham. Improved function was associated with mixed exercise vs. usual care, and Pilates vs. no intervention. Temporary increased minor pain was associated with mixed exercise vs. no intervention, and yoga vs. usual care. Little to no difference was found for other comparisons and outcomes. When pooling exercise types, exercise vs. no intervention probably reduces pain in adults (8 RCTs, SMD = − 0.33, 95% CI − 0.58 to − 0.08) and functional limitations in adults and older adults (8 RCTs, SMD = − 0.31, 95% CI − 0.57 to − 0.05) (moderate certainty evidence).

Conclusions With moderate certainty, structured exercise programs probably reduce pain and functional limitations in adults and older people with CPLBP.

Keywords Low back pain · Systematic review · Meta-analysis · Exercise · Rehabilitation

Introduction

Exercise therapy or structured exercise programs are widely used to manage low back pain (LBP). Exercise therapy is defined as "a series of specific movements with the aim of training or developing the body by a routine practice or as physical training to promote good physical health" [1] with a goal to reduce pain and functional limitations. Exercise therapies are prescribed or planned by health practitioners and include conducting postures, movements, and/or activities (e.g., strengthening, stretching, aerobic exercise)

at varying dosages (duration, frequency, intensity) [2]. For people with chronic primary LBP (CPLBP), exercise therapy may improve musculoskeletal function, while also benefiting most other body systems and mental wellbeing [3]. In turn, this may reduce pain and functional limitations, and improve emotional and psychological wellbeing [2]. Exercise therapy is accessible globally.

Hayden and colleagues published a Cochrane review (2021) (literature search date ending 28 April 2018) to assess the impact of exercise therapy on pain and functional limitations for the management of chronic LBP in adults compared to placebo, no treatment, or usual care (pooled together), or other conservative treatments (249 randomized controlled trials (RCTs); 24,486 participants) [2] and a

---

Extended author information available on the last page of the article

network meta-analysis comparing different types of exercise treatments [4]. They concluded with moderate certainty that exercise reduces pain and functional limitations when compared to no treatment, usual care, or sham, but not when compared to other conservative treatments [2].

To develop clinical practice guideline recommendations for the management of CPLBP in adults, the WHO commissioned the current systematic review to update the evidence and expand the aims of Hayden et al.'s previously published Cochrane review [2] by assessing additional important outcomes, conducting additional subgroup analyses, and disaggregating pairwise findings by exercise type (compared to no treatment, placebo/sham, or usual care).

The objectives of this systematic review of RCTs were to determine: (1) the benefits and harms of structured exercise programs compared to placebo/sham, usual care, or no intervention for the management of CPLBP in adults, including older adults (aged ≥ 60 years); and (2) whether the benefits and harms of structured exercise programs vary by age, gender/sex, presence of leg pain, race/ethnicity, or national economic development of the countries where the RCTs were conducted.

Methods

This systematic review was conducted as part of a series of reviews to inform a WHO clinical practice guideline on the management of CPLBP in adults. The development of this guideline was ongoing at the time of submission of this manuscript. The review was conducted in collaboration with the Cochrane 'exercise treatment for chronic low back pain' collaborative review team, led by Prof. Jill Hayden [5]. The methods are detailed in the methodology article of this series [6].

Briefly, we updated and expanded the scope of the previously published Cochrane review [2]. The current review differs from Hayden et al.'s in the following ways: 1) we updated the literature search to include RCTs published from 28 April 2018 through 17 May 2022; 2) we assessed additional outcomes identified as critical by the WHO Guideline Development Group (GDG); 3) we conducted additional subgroup analyses (e.g., age, gender/sex); 4) we analyzed and reported the results separately for different exercise types, specifically comparing the effects of each exercise intervention to its respective comparator; 5) we did not assess 'other conservative treatment' comparisons (e.g., exercise vs. manual therapy); 6) we excluded RCTs of multimodal interventions where the specific effects of exercise could not be isolated; 7) we excluded RCTs judged to have high risk of bias in our primary analyses (although included all RCTs, irrespective

of risk of bias in a supplementary analysis); and 8) the eligibility criteria for the population of interest differed to some degree. For example, we did not exclude RCTs of participants who had specific pathologies (e.g., disc herniation, lumbar spinal stenosis, and spondylolisthesis) provided all other eligibility criteria were satisfied. We also did not exclude RCTs of surgical populations if time since surgery was at least 12 months and participants had no history of fusion and/or disc replacement surgery.

We registered our review protocol with PROSPERO (International Prospective Register of Systematic Reviews) (CRD42022314576) on 7 March 2022.

In collaboration with the Cochrane review team, we modified the original search strategy using a detailed search optimization process [7]. The updated strategy was approved by a Cochrane musculoskeletal (MSK) literature search specialist. We searched MEDLINE (Ovid), CENTRAL (Cochrane Library, Wiley), and Embase (Elsevier) with no date or language restrictions up until 17 May 2022 (see Online Resource 1). Retrieved citations were de-duplicated against the search results of the previous Cochrane review update.

We included RCTs that compared structured exercise programs to placebo/sham, usual care, and no intervention (including comparison interventions where the attributable effect of exercise could be isolated, i.e., exercise + medication vs. same medication alone) in adults (aged ≥ 20 years) with CPLBP. Eligible interventions included all types of exercise with no exclusions based on setting, mode of delivery (e.g., in-person vs. telehealth, group vs. individual, home vs. clinic or community) or degree of personalization (standardized vs. individualized). Individuals may have been given verbal or written exercise instructions (e.g., handbook). Eligible exercise interventions, considered as separate exercise types, included, but were not limited to aerobic exercise; muscle strength training; stretching, flexibility or mobilizing exercises; yoga; core strengthening; motor control exercise; functional restoration exercise (not including multimodal programs of exercise with other interventions, such as psychological supports); Pilates; Tai Chi; Qigong; and mixed exercise therapies (i.e., two or more types of exercise in which one did not clearly predominate).

In addition to the main critical outcomes assessed for all reviews in this series (pain, function, health-related quality of life (HRQoL), harms, psychological functioning, and social participation including work), we also assessed additional critical outcomes requested by the WHO GDG for this review – the change in use of medications, burden related to the intervention or comparator (e.g., ease of access to the intervention, time burden of the intervention), performance-based physical functioning, and falls (older adults only aged ≥ 60 years). We reported

outcomes based on post-intervention follow-up intervals including: (1) immediate term (closest to 2 weeks after the intervention period); (2) short term (closest to 3 months after the intervention period); (3) intermediate term (closest to 6 months after the intervention period); (4) long term (closest to 12 months after the intervention period); and (5) extra-long term (more than 12 months after the intervention period).

We assessed between-group differences to determine the magnitude of the effect of an intervention and to assess its effectiveness [8, 9] (details in the methodology article in this series) [6]. Briefly, we considered a mean difference (MD) of ≥ 10% of the scale range or ≥ 10% difference in risk for dichotomous outcomes to be a minimally important difference (MID) [10, 11]. If the standardized mean difference (SMD) was calculated, SMD ≥ 0.2 was considered a MID [12].

Pairs of reviewers independently screened studies for eligibility, and critically appraised risk of bias (ROB) using the Cochrane ROB 1 tool [13], modified from the Cochrane Back and Neck Methods Guidelines [14]. One reviewer extracted data for all included RCTs, which was then verified by a second reviewer. Any disagreements were resolved by consensus between paired reviewers or with a third reviewer, when necessary. Forms and guidance for screening, risk of bias assessment, and data extraction were adapted from those developed by Hayden et al. in the conduct of the 'exercise for chronic low back pain' collaborative review, in which members of our team participated [5]. The forms were completed using DistillerSR Inc. [15]—a web-based electronic systematic review software application.

In our primary synthesis, our analyses were conducted according to exercise type (e.g., aerobic exercise, yoga). In addition to the subgroup analyses conducted for all reviews in this series (age, gender/sex, presence of leg pain, race/ethnicity, and national economic development of country where RCT was conducted), we aimed to perform subgroup analyses according to exercise dosage and intensity, and to conduct a sensitivity analysis by removing RCTs rated as unclear ROB.

We conducted random-effects meta-analyses and narrative synthesis where meta-analysis was not appropriate [16], and graded the certainty of evidence using Grading of Recommendations Assessment, Development and Evaluation (GRADE) [17]. The comparisons involving no intervention and interventions where the attributable effect of exercise could be isolated were combined in meta-analyses. Meta-analyses were conducted using R statistical packages [18, 19], and GRADE Evidence Profiles and GRADE Summary of Findings tables were developed using GRADEpro software [20].

Following completion of our primary synthesis, the WHO commissioned a supplementary evidence synthesis to further inform the formulation of recommendations by the GDG. In the supplementary evidence synthesis, we synthesized the 13 RCTs (judged as low or unclear ROB) included in our primary evidence synthesis along with 55 additional RCTs originally excluded from our synthesis due to high ROB. These studies were identified as having been published in the period 28 April 2018 (search end date of Hayden's previously published Cochrane review [2]) to 17 May 2022. We included all 13 trials from the primary synthesis (from database inception through 17 May 2022) in this supplementary synthesis since no differences in the magnitude or directions of the effect estimates were observed in a sensitivity analysis where RCTs published on or before 28 April 2018 were excluded.

In the supplementary evidence synthesis (see Online Resource 8), we included RCTs that compared *any structured exercise program or exercise type* to the same comparisons as in our primary synthesis. The outcomes assessed were pain, function, and harms only. The key differences between the primary and supplementary evidence syntheses are summarized (Table 1).

The WHO was provided with the primary and supplementary evidence syntheses to support the GDG in formulation of recommendations. The GDG may have also considered other aligned evidence when formulating its recommendations (currently under development).

Results

Our electronic search strategy identified 8592 new citations (Fig. 1), the Cochrane RCT Classifier/known assessments and Cochrane Crowd first excluded 6131 non-RCTs (RCT Classifier/known assessments: 3281, Cochrane Crowd: 2850). We subsequently screened 2503 records and 398 full-text reports. Of these, 69 new RCTs were eligible. We included an additional 55 RCTs from the published review [2], which totalled 124 RCTs. Of these, 111 were excluded from the primary analyses due to an overall high ROB rating (see Online Resource 2). Therefore, we included 13 RCTs ($n = 1362$ participants) in our synthesis [21–33] ranging from 45 to 313 participants per trial, predominantly from healthcare settings (see Online Resources 3, 4).

Regarding unpublished RCTs, we identified 185 RCTs (registrations and published protocols) in the WHO ICTRP. Of these, 14 authors could not be contacted because an email address could not be located. Thus, 171 authors were contacted and 164 received our invitation to respond to a REDCap survey [34, 35] consisting of our specific queries. Of these, 32% (53/164) responded; 19 reported that their RCT would not meet our inclusion criteria; 26 reported their RCT was ongoing; and 8 provided citations, which we

Table 1 Differences between the primary and supplementary evidence syntheses

Evidence synthesis component	Primary evidence synthesis	Supplementary evidence synthesis
Search period	Database inception through 17 May 2022 (for low or unclear ROB RCTs)	Database inception through 17 May 2022 (for low or unclear ROB RCTs)+28 April 2018 through 17 May 2022 for high ROB RCTs[a]
Inclusion criteria based on ROB	Low or unclear ROB	Low, unclear or high ROB
Outcomes	Pain, function, harms, health-related quality of life, psychological functioning, social participation (+change in use of medications, burden related to treatment, performance-based physical functioning, and falls in adults aged ≥ 60 years)	Pain, function, harms
Sub-group comparisons	Age, national economic development, exercise type, ROB judgement (low vs. unclear), gender/sex, presence of leg pain, race/ethnicity	Age, national economic development, exercise type, ROB judgement (low vs. not low)[b]

ROB: risk of bias

[a]Randomized controlled trials (RCTs) published on or before 28 April 2018 were included in the supplementary synthesis since no differences in the magnitude or direction of effect estimates were observed in sensitivity analyses when these RCTs were excluded

[b]Other subgroups were not analyzed (i.e., gender/sex, presence of leg pain, race/ethnicity) as the primary synthesis did not demonstrate varied findings
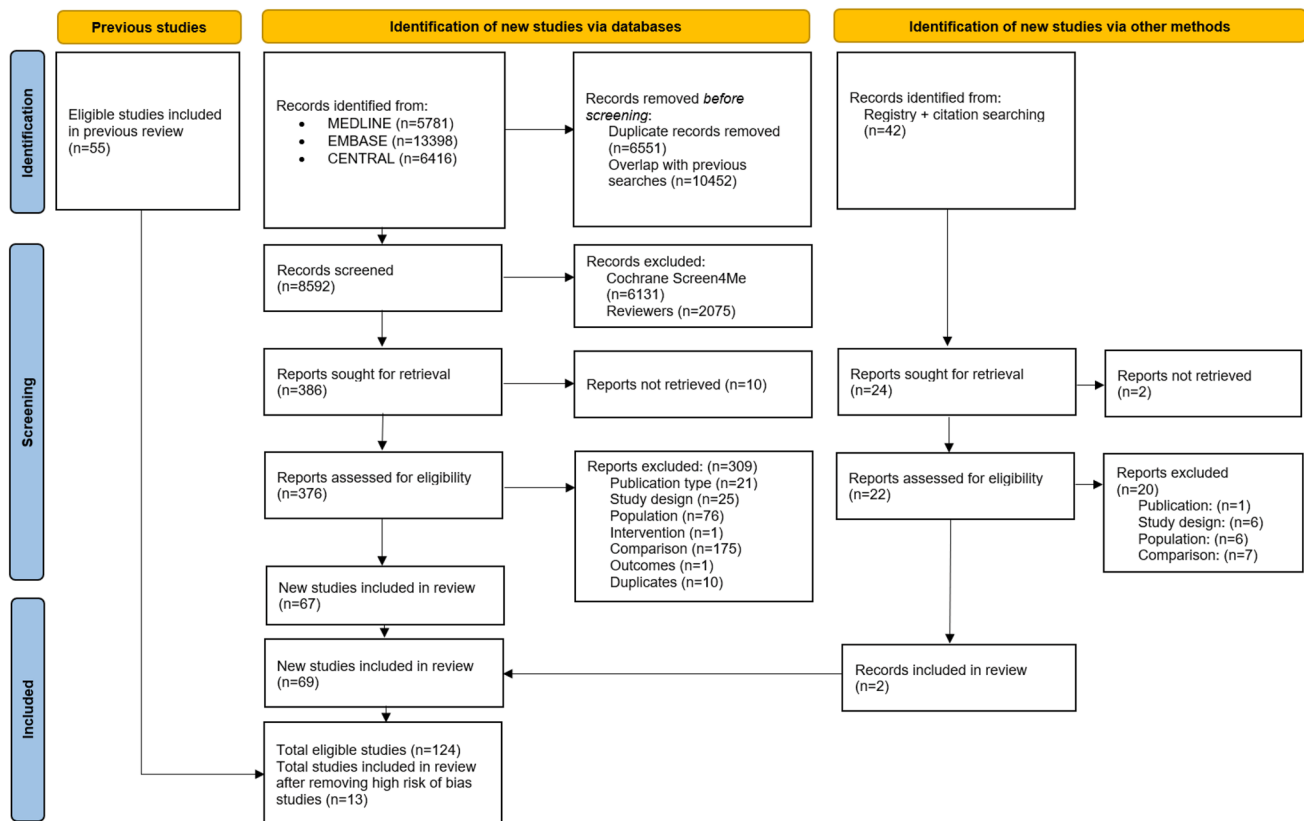

Fig. 1 Flow diagram of literature search

confirmed were already included in our review. Thus, we did not include any unpublished RCTs in our review.

The 13 included RCTs were conducted in high-income economies [36]: Australia (1 RCT) [22], Germany (1 RCT) [28], Japan (1 RCT) [24], Netherlands (1 RCT) [30], United Kingdom (1 RCT) [32], and the United States (1 RCT) [33]; upper-middle income economies: Brazil (3 RCTs) [23, 25, 26] and Turkey (1 RCT) [31]; and lower-middle income

economies: India (1 RCT) [21] and Iran (2 RCTs) [27, 29]. The mean age ranged from 37 to 74 years; two RCTs assessed older adults ($n=252$) [24, 33]. The percentage of females within the RCTs ranged from 27 to 84%. None of the RCTs included adults who all had leg pain in addition to back pain. Two RCTs reported on adults without leg pain [23, 29]. In one RCT, adults had CPLBP either with or without non-radicular leg pain [30]; and in another adults had CPLBP either with or without unclassified (radicular vs. non-radicular) leg pain [33]. The presence of leg pain was not classified in nine RCTs [21–28, 31, 32]. Where reported by authors, the mean duration of CPLBP ranged from 7 months to ≥ 15 years.

The RCTs assessed aerobic exercise [26, 28]; core strengthening exercise [27]; muscle strength training [23, 29]; mixed exercise [21, 24, 30, 31, 33]; Pilates [25]; stretching, flexibility or mobilizing exercises [29]; yoga [32]; and motor control exercise [22]. We did not identify any low or unclear ROB trials that assessed the other exercise types: functional restoration, Tai Chi, Qigong, or hydrotherapy/aquatic exercise. The number of exercise sessions delivered ranged from 8 to 30, with the duration of each session ranging from 15 to 105 min. Exercise was compared to interventions where the attributable effects of exercise could be isolated, sham, or usual care, and most RCTs assessed pain and function in the immediate term (Table 2). The outcomes were assessed in the immediate term (13 RCTs) [21–33], short term (3 RCTs) [24, 28, 31], intermediate term (8 RCTs) [22–25, 30–33], and long term (4 RCTs) [22, 23, 30, 32] (Table 2). The RCTs were rated as overall unclear ROB (11, 85%), or low ROB (2, 15%) (Online Resource 2). The agreement on ROB judgements was high (weighted overall kappa score 0.92).

Certainty of Evidence

The certainty of the evidence ranged from very low (for outcomes assessed with the individual exercise types) to moderate (for outcomes assessed after pooling exercise types). Certainty of evidence was downgraded due to ROB, inconsistency, indirectness, and/or imprecision of the effect estimates (see Online Resources 5, 6 and 7). For results reported as a MD, lower or negative values refer to reduced pain, functional limitations, depression, or fear avoidance; higher or positive values refer to improved HRQoL and self-efficacy.

Aerobic Exercise Versus Comparison Interventions With Isolated Exercise Effects

The certainty of evidence was very low for all outcomes. It is uncertain whether aerobic exercise reduces *pain (scale 0 to 10, 0 = no pain)* in the immediate (2 RCTs; MD = − 1.33, 95% confidence interval (CI) − 2.27 to − 0.40) (plot 1.1.1.1) [26, 28], or short term (1 RCT; MD = − 1.26, 95% CI − 2.51 to − 0.01) (plot 1.1.1.2) [28]. It is uncertain whether aerobic exercise makes little or no difference to *functional limitations (scale 0 to 100, 0 = no functional limitations)* in the immediate (2 RCTs; MD = − 1.30, 95% CI − 3.89 to 1.29) (plot 1.1.2.1), [26, 28] or short term (1 RCT; MD = 0.90, 95% CI − 5.66 to 7.46) (plot 1.1.2.2) [28]. It is uncertain whether aerobic exercise makes little or no difference to *HRQoL (scale 0 to 100, 0 = poor quality of life; PCS = physical component summary; MCS = mental component summary)* in the immediate and short terms (immediate: *PCS:* MD = 3.50, 95% CI − 0.05 to 7.05; *MCS:* MD = − 1.20, 95% CI − 5.22 to 2.82; plot 1.1.3.1.1 and 1.1.3.1.2; short term: *PCS:* MD = 3.70, 95% CI 0.05 to 7.35; *MCS:* MD = 2.20, 95% CI − 3.15 to 7.55; plot 1.1.3.2.1 and 1.1.3.2.2) [28]. It is uncertain whether aerobic exercise makes little or no difference to *adverse events/harms* (1 RCT) (no plot, narrative synthesis). Authors reported no adverse events [28].

Core Strengthening Exercise Versus Comparison Interventions With Isolated Exercise Effects

The certainty of evidence was very low for all outcomes and based on one RCT [27]. In the immediate term, it is uncertain whether core strengthening reduces *pain (scale 0 to 10, 0 = no pain)* (MD = − 0.56, 95% CI − 0.94 to − 0.19) (plot 2.1.1.1), or *functional limitations (scale 0 to 24, 0 = no functional limitations)* (MD = − 1.7, 95% CI − 2.42 to − 0.98) (plot 2.1.2.1). It is uncertain whether core strengthening exercise makes little or no difference to *adverse events/harms* (no plot, narrative synthesis).

Muscle Strength Training Versus Comparison Interventions With Isolated Exercise Effects

The certainty of evidence was very low for all outcomes. It is uncertain whether muscle strength training makes little or no difference to *pain (scale 0 to 10, 0 = no pain)* in the immediate (2 RCTs; MD = − 0.39, 95% CI − 1.16 to 0.38) (plot 3.1.1.1) [23, 29], intermediate (1 RCT; MD = − 0.40, 95% CI − 1.67 to 0.87) (plot, 3.1.1.2) [23], or long term (1 RCT; MD = − 0.10, 95% CI − 1.32 to 1.12) (plot, 3.1.1.3) [23]. It is uncertain whether muscle strength training makes little or no difference to *function (benefit indicated by lower values)* in the immediate (2 RCTs; standardized mean difference (SMD) = 0.05, 95% CI − 0.34 to 0.45) (plot 3.1.2.1) [23, 29]; or intermediate (1 RCT; MD = − 0.60, 95% CI − 3.20 to 2.00) (plot 3.1.2.2) [23], and long terms (1 RCT; MD = − 0.20, 95% CI − 2.73 to 2.33) (plot 3.1.2.3) (*scale 0 to 24, 0 = no functional limitations*) [23].

Table 2 Number of included RCTs by comparison and outcome

Outcome assessed (post-intervention)	Follow-up			
	Immediate (2 weeks)	Short (3 months)	Intermediate (6 months)	Long (12 months)
Exercise versus no intervention[a] (9 RCTs[b], 10 exercise groups)				
Pain	**10[c]**	*2*	**5[c]**	*3*
Function	**10[c]**	*2*	**5[c]**	*2*
Health-related quality of life	**3[c]**	*2*	*2[c]*	*1*
Psychological functioning	**3[c]**	–	**3[c]**	*1*
Social participation	–	–	–	–
Change in medication use	–	–	–	–
Treatment-related burden	–	–	–	–
Performance-based physical functioning	*1[c]*	–	*1[c]*	–
Falls	–	–	–	–
Harms	**5[c]**			
Number of RCTs per exercise type: aerobic = 2, core strengthening = 1, general strength training = 2, mixed exercise = 3[c], Pilates = 1, stretching/flexibility/mobilizing exercise = 1				
Exercise versus sham				
Pain	*1*	–	*1*	*1*
Function	*1*	–	*1*	*1*
Health-related quality of life	–	–	–	–
Psychological functioning	–	–	–	–
Social participation	–	–	–	–
Change in medication use	–	–	–	–
Treatment-related burden	–	–	–	–
Performance-based physical functioning	–	–	–	–
Falls	–	–	–	–
Harms	1			
Number of RCTs per exercise type: motor control = 1				
Exercise versus usual care				
Pain	**3[c]**	*1[c]*	*2[c]*	*1*
Function	**3[c]**	*1[c]*	*2[c]*	*1*
Health-related quality of life	*2[c]*	*1[c]*	*2[c]*	*1*
Psychological functioning	*2[c]*	*1[c]*	*2[c]*	*1*
Social participation	–	–	–	–
Change in medication use	–	–	–	–
Treatment-related burden	–	–	–	–
Performance-based physical functioning	–	–	–	–
Falls	–	–	–	–
Harms	1			
Number of RCTs per exercise type: mixed exercise = 2[c], yoga = 1				

Bold values: majority of studies are in this category, italic values: some studies

[a]Included comparison interventions where the attributable effect of exercise could be isolated (i.e., combined exercise with treatment B versus treatment B alone)

[b]One RCT reported two intervention groups: 1) hamstring static stretching + physiotherapy vs. physiotherapy, 2) hamstring strengthening in lengthened position + physiotherapy vs. physiotherapy

[c]One RCT included adults aged ≥ 60 years

Mixed Exercise Versus Comparison Interventions With Isolated Exercise Effects

All Adults

Due to very low certainty evidence, it is uncertain whether mixed exercise makes little or no difference to **pain** in the immediate (2 RCTs; SMD = − 0.01, 95% CI − 0.32 to 0.31; *benefit indicated by lower values*) (plot 4.1.1.1) [31, 33], short (1 RCT; MD = − 0.10, 95% CI − 1.34 to 1.14; *scale 0 to 10, 0 = no pain*) (plot 4.1.1.2) [31], intermediate (2 RCTs; SMD = 0.03, 95% CI − 0.23 to 0.29; *benefit indicated by lower values*) (plot 4.1.1.3) [31, 33], or long term (1 RCT; MD = 8.88, 95% CI − 0.36 to 18.13; *scale 0 to 100, 0 = no pain*) (no plot, narrative synthesis) [30]. Mixed exercise may make little or no difference to **function** in the immediate term (2 RCTs; SMD = − 0.15, 95% CI − 0.48 to 0.18; *benefit indicated by lower values*; low certainty evidence) (plot 4.1.2.1) [31, 33]. Due to very low certainty evidence, it is uncertain whether mixed exercise makes little or no difference to function in the short (1 trial; MD = − 1.25, 95% CI − 2.79 to 0.29; *scale 0 to 9, 0 = no functional limitations*) (plot 4.1.2.1) [31], intermediate (2 trials; SMD = − 0.09, 95% CI − 0.42 to 0.24; *benefit indicated by lower values*) (plot 4.1.2.2) [31, 33], or long term (1 trial; MD = 1.62, 95% CI − 0.06 to 3.31; *scale 0 to 24, 0 = no functional limitations*) (no plot, narrative synthesis) [30]. Due to very low certainty evidence from one RCT [31], it is uncertain whether mixed exercise makes little or no difference to **HRQoL** *(scale 0 to 3, 0 = poor quality of life)* in the immediate (MD = 0.24, 95% CI − 0.06 to 0.54) (plot 4.1.3.1), short (MD = 0.17, 95% CI − 0.07 to 0.41) (plot 4.1.3.2), or intermediate term (MD = 0.19, 95% CI − 0.09 to 0.47) (plot 4.1.3.3); or **depression** *(scale 0 to 63, 0 = no depression)* in the long term (MD = − 0.09, 95% CI − 2.11 to 1.93) (no plot, narrative synthesis). It is uncertain whether mixed exercise makes little or no difference to **adverse events/harms** (2 RCTs; odds ratio (OR) 4.24, 95% CI 0.69 to 25.95; very low certainty evidence) (plot 4.1.5) [30, 33]. Adverse events were mainly minor and included back and knee pain.

Older Adults

Due to very low certainty evidence from 1 RCT [33], in older adults, it is uncertain if mixed exercise makes little or no difference to **pain** *(benefit indicated by lower values)* in the immediate (SMD = − 0.10, 95% CI − 0.44 to 0.23) (plot 4.1.6.1.1), or intermediate term (SMD = − 0.01, 95% CI − 0.39 to 0.40) (plot 4.1.6.1.2); **function** *(benefit indicated by lower values)* in the immediate (SMD = − 0.01,

95% CI − 0.29 to 0.27) (plot 4.1.6.2.1), or intermediate term (SMD = 0.03, 95% CI − 0.24 to 0.31) (plot 4.1.6.2.2); **depression** *(scale 0 to 30, 0 = no depression)* in the immediate (MD = − 0.11, 95% CI − 1.87 to 1.66) (plot 4.1.4.1) or intermediate term (MD = 0.14, 95% CI − 1.92 to 2.20) (plot 4.1.4.2); **self-efficacy** *(scale 10–100, benefit indicated by higher values)* in the immediate (between-group difference change score = 2.1, standard error (SE) 3.1, *p* = 0.50), or intermediate term (between-group difference change score = − 0.8, SE 3.2, *p* = 0.80) (narrative synthesis); **catastrophizing and fear avoidance** *(benefit indicated by lower values)* in the immediate and intermediate terms (no plots, narrative synthesis); **performance-based physical functioning** in the immediate (between-group difference change scores: *usual pace gait speed:* 0.02 m/second, *p* = 0.29; *chair raise time:* − 0.8 s, *p* = 0.008; *stair climb time*: − 0.0 s, *p* = 0.99) or intermediate term (between-group difference change scores: *usual pace gait speed:* 0.00 m/second, *p* = 0.92; *chair raise time:* 0.1 s, *p* = 0.88; *stair climb time:* − 0.6 s, *p* = 0.61) (no plots, narrative synthesis); or **harms** (OR = 3.06, 95% CI 0.31 to 29.93) (plot 4.1.6.3). One participant experienced increased back pain. Authors reported no substantial intervention-associated adverse events.

Due to very low certainty evidence from 1 RCT [33], in older adults in the immediate term, it is uncertain whether mixed exercise worsens **HRQoL PCS** *(scale 0 to 100, 0 = poor quality of life)* (MD = − 6.56, 95% CI − 13.03 to − 0.10) (plot 4.1.3.3.1). Mixed exercise may make little or no difference to **HRQoL MCS** in the immediate (MD = − 1.05, 95% CI − 4.38 to 2.28) (plot 4.1.3.1.2); or intermediate term (**PCS:** MD = − 2.31, 95% CI − 9.33 to 4.70; **MCS:** MD = − 0.83, 95% CI − 8.67 to 7.00) (plot 4.1.3.3.2).

Mixed Exercise vs. Usual Care

All Adults

For outcomes that are based on RCTs of older adults only, results are reported under *older adults* below.

Due to very low certainty evidence, it is uncertain whether mixed exercise makes little or no difference to **pain** *(scale 0 to 10, 0 = no pain)* in the immediate (2 RCTs; MD = − 0.12, 95% CI − 0.91 to 0.68) (plot 4.2.1.1) [21, 24], short (1 RCT; MD = − 0.30, 95% CI − 1.66 to 1.06) (plot 4.2.1.2) [24], or intermediate term (1 RCT; MD = 0.00, 95% CI − 1.26 to 1.26) (plot 4.2.1.3) [24]. It is uncertain whether mixed exercise reduces **functional limitations** *(benefit indicated by lower values)* in the immediate term (2 RCTs; SMD = − 0.62, 95% CI − 0.96 to − 0.28) (plot 4.2.2.1) [21, 24].

Older Adults

Due to very low certainty evidence from one RCT of older adults [24], it is uncertain whether mixed exercise makes little or no difference to **pain** *(scale 0 to 10, 0 = no pain)* in the immediate (MD = − 0.80, 95% CI − 2.42 to 0.82) (plot 4.2.1.1), short (MD = − 0.30, 95% CI − 1.66 to 1.06) (plot 4.2.1.2), or intermediate term (MD = 0.00, 95% CI − 1.26 to 1.26) (plot 4.2.1.3); **HRQoL** *(scale 0 to 1, 0 = poor quality of life)* in the immediate (MD = 0.05, 95% CI − 0.01 to 0.11) (plot 4.2.3.1), short (MD = 0.04, 95% CI − 0.00 to 0.08) (plot 4.2.3.2), or intermediate term (MD = 0.05, 95% CI − 0.00 to 0.10) (plot 4.2.3.3); or **self-efficacy** *(scale 0 to 60, 0 = poor self-efficacy)* in the immediate (MD = 3.00, 95% CI − 2.39 to 8.39) (plot 4.2.4.1), short (MD = 3.00, 95% CI − 1.63 to 7.63) (plot 4.2.4.2), or intermediate term (MD = 4.00, 95% CI − 3.81 to 11.81) (plot 4.2.4.3). It is uncertain whether mixed exercise reduces **functional limitations** in the immediate term (SMD = − 0.86, 95% CI − 1.45 to − 0.27; *benefit indicated by lower values*) (plot 4.2.2.1). It is uncertain whether mixed exercise makes little or no difference to functional limitations *(scale 0 to 24, 0 = no functional limitations)* in the short (MD = − 2.30, 95% CI − 4.92 to 0.32) (plot 4.2.2.2), or intermediate term (MD = − 2.50, 95% CI − 5.19 to 0.19) (plot 4.2.2.3).

Pilates Exercises Versus Comparison Interventions With Isolated Exercise Effects

Due to very low certainty evidence from one RCT [25], it is uncertain whether Pilates reduces **pain** *(scale 0 to 10, 0 = no pain)* in the immediate term (MD = − 2.10, 95% CI − 3.07 to − 1.13) (plot 5.1.1.1), or makes little or no difference to pain in the intermediate term (MD = − 0.80, 95% CI − 1.75 to 0.15) (plot 5.1.1.2). It is uncertain whether Pilates reduces **functional limitations** *(scale 0 to 24, 0 = no disability)* in the immediate (MD = − 3.50, 95% CI − 5.48 to − 1.52) (plot 5.1.2.1), or intermediate term (MD = − 2.20, 95% CI − 4.35 to − 0.05) (plot 5.1.2.2).

Due to very low certainty evidence from one RCT [25], it is uncertain whether Pilates makes little or no difference to **fear avoidance** *(scale 17–68, benefit indicated by lower values)* in the immediate (MD = − 1.80, 95% CI − 5.12 to 1.52) (plot 5.1.3.1), or intermediate term (MD = − 0.80, 95% CI − 3.86 to 2.26) (plot 5.1.3.2); or to **harms**: authors reported no adverse events (no plot, narrative synthesis).

Stretching, Flexibility Or Mobilizing Exercises Versus Comparison Interventions With Isolated Exercise Effects

Due to very low certainty evidence from one RCT [29], in the immediate term, it is uncertain whether stretching, flexibility or mobilizing exercise makes little or no difference to **pain** *(scale 0 to 10, 0 = no pain)* (MD = − 0.18, 95% CI − 1.61 to 1.25) (plot 6.1.1.1) or **function** *(scale 0 to 100, 0 = no disability)* (MD = − 3.97, 95% CI − 13.14 to 5.19) (plot 6.1.2.1).

Yoga Versus Usual Care

The evidence is based on one RCT [32] and is very low certainty for all outcomes and time points. The results in this section are narratively synthesized (no forest plots).

It is uncertain whether yoga makes little or no difference to **pain** *(scale 0 to 100, 0 = no pain)* in the immediate (between-group difference in means = − 2.42, 95% CI − 4.97 to 0.12), intermediate (between-group difference in means = − 1.74, 95% CI − 4.32 to 0.84), or long term (between-group difference in means = − 0.73, 95% CI − 3.30 to 1.84); or **HRQoL** *(scale 0 to 100, 0 = poor quality of life)* in the immediate (between-group difference in means: **PCS:** 1.36, 95% CI − 0.70 to 3.41; **MCS:** 2.02, 95% CI − 0.31 to 4.35), intermediate (between-group difference in means: **PCS:** 1.24, 95% CI − 0.83 to 3.33; **MCS:** 2.02, 95% CI − 0.34 to 4.37), or long term (between-group difference in means: **PCS:** 0.80, 95% CI − 1.28 to 2.87; **MCS:** 0.42, 95% CI − 1.92 to 2.77). It is uncertain whether yoga reduces **functional limitations** *(scale 0 to 24, 0 = no disability)* in the immediate (between-group difference in means = − 2.17, 95% CI − 3.31 to − 1.03), intermediate (between-group difference in means = − 1.48, 95% CI − 2.62 to − 0.03), or long term (between-group difference in means = − 1.57, 95% CI − 2.71 to − 0.42). It is uncertain whether yoga improves **self-efficacy** *(scale 0 to 60, 0 = poor self-efficacy)* in the immediate (between-group difference in means = 2.96, 95% CI 0.35 to 5.58), or intermediate term (between-group difference in means = 3.33, 95% CI 0.68 to 5.97); or whether yoga makes little or no difference to self-efficacy in the long term (between-group difference in means = 1.75, 95% CI − 0.87 to 4.38). It is uncertain whether yoga increases **minor adverse events/harms** (i.e., increased pain) (OR 25.77, 95% CI 1.50 to 441.85) (plot 7.1.1); or whether yoga makes little or no difference to **serious adverse events/harms** (OR 0.51, 95% CI 0.05 to 5.70) (plot 7.1.2). Authors reported one participant in the yoga group experienced severe pain (typically does after physical activity).

Motor Control Exercise Versus Sham

The evidence is based on one RCT [22] and is very low certainty for all outcomes. It is uncertain whether motor control exercise reduces **pain** *(scale 0 to 10, 0 = no pain)* in the immediate (MD = − 1.00, 95% CI − 1.85 to − 0.15) (plot 8.1.1.1) or long term (MD = − 1.30, 95% CI − 2.13 to − 0.47) (plot 8.1.1.3), or whether it makes little or no

difference to pain in the intermediate term (MD = − 0.60, 95% CI − 1.46 to 0.26) (plot 8.1.1.2). It is uncertain whether motor control exercise reduces *functional limitations (scale 0 to 24, 0 = no disability)* in the immediate term (MD = − 2.30, 95% CI − 4.26 to − 0.34) (plot 8.1.2.1); or whether it makes little or no difference in the intermediate (MD = − 1.90, 95% CI − 4.06 to 0.26) (plot 8.1.2.2), or long term (MD = − 0.90, 95% CI − 3.15 to 1.35) (plot 8.1.2.3). It is uncertain whether motor control exercise makes little or no difference to *harms* (OR = 1.52, 95% CI 0.25 to 9.36) (plot 8.1.3). Authors reported all adverse events were temporary exacerbations of pain.

Pooled Analysis of All Exercise Types Versus Comparison Interventions With Isolated Exercise Effects

We conducted a post hoc analysis by pooling all exercise types since only 1–3 RCTs were identified for each exercise type and none on their own showed a clear benefit. To be included in this analysis, data from two or more of the eight exercise types had to be available per comparison, outcome, and time point. Otherwise, findings of the individual eight

exercise types have been reported in the eight previous comparisons.

Exercise probably reduces *pain (benefit indicated by lower values)* in the immediate term (8 RCTs; SMD = − 0.33, 95% CI − 0.58 to − 0.08; moderate certainty evidence) (Fig. 2) (plot 9.1.1.1) [23, 25–29, 31, 33]. Due to very low certainty evidence, it is uncertain whether exercise makes little or no difference to *pain (scale 0 to 10, 0 = no pain)* in the short (2 RCTs; MD = − 0.68, 95% CI − 1.82 to 0.46) (plot 9.1.1.2) [28, 31], or long term (1 RCT; between-group MD = 8.88, 95% CI − 0.36 to 18.13; scale 0 to 100, 0 = no pain) (no plot, narrative synthesis) [30]. Exercise may make little or no difference to *pain (benefit indicated by lower values)* in the intermediate term (4 RCTs; SMD = − 0.08, 95% CI − 0.29 to 0.13; low certainty evidence) (plot 9.1.1.3) [23, 25, 31, 33].

Exercise probably reduces *functional limitations (benefit indicated by lower values)* in the immediate term (8 RCTs; SMD = − 0.31, 95% CI − 0.57 to − 0.05; moderate certainty evidence) (Fig. 3) (plot 9.1.2.1) [23, 25–29, 31, 33]. Due to very low certainty evidence, it is uncertain whether exercise makes little or no difference to function in the short (2 RCTs; SMD = − 0.26, 95% CI − 0.67 to 0.14; *benefit indicated by*

Study or Subgroup	Exercise Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Fukuda 2021	2.3	2.2	35	2.9	2	35	11.4%	-0.28 [-0.75, 0.19]
Miyamoto 2013	3.1	2.3	43	5.2	2.3	43	11.9%	-0.90 [-1.35, -0.46]
Nardin 2022	1.6	2	20	3.5	2.7	20	8.4%	-0.78 [-1.43, -0.14]
Rahbar 2018	16.22	8.66	40	21.84	8.54	40	11.8%	-0.65 [-1.10, -0.20]
Rotter 2022	35.9	21.19	24	45.4	21.55	24	9.5%	-0.44 [-1.01, 0.14]
Shamsi 2022a (same study as Shamsi 2022b)	25.38	14.22	15	26.07	19.43	8	5.9%	-0.04 [-0.90, 0.82]
Shamsi 2022b (same study as Shamsi 2022a)	24.26	20.57	15	26.07	19.43	7	5.5%	-0.09 [-0.98, 0.81]
Takinaci 2019	44.27	27.14	26	36.83	21.3	24	9.8%	0.30 [-0.26, 0.86]
Weiner 2008a (same study as Weiner 2008b)	8.1	8.2	50	10.5	9.2	50	12.9%	-0.27 [-0.67, 0.12]
Weiner 2008b (same study as Weiner 2008a)	8.9	7.9	50	8.4	6.3	50	13.0%	0.07 [-0.32, 0.46]
Total (95% CI)			318			301	100.0%	-0.33 [-0.58, -0.08]

Heterogeneity: Tau² = 0.08; Chi² = 20.05, df = 9 (P = 0.02); I² = 55%
Test for overall effect: Z = 2.57 (P = 0.01)

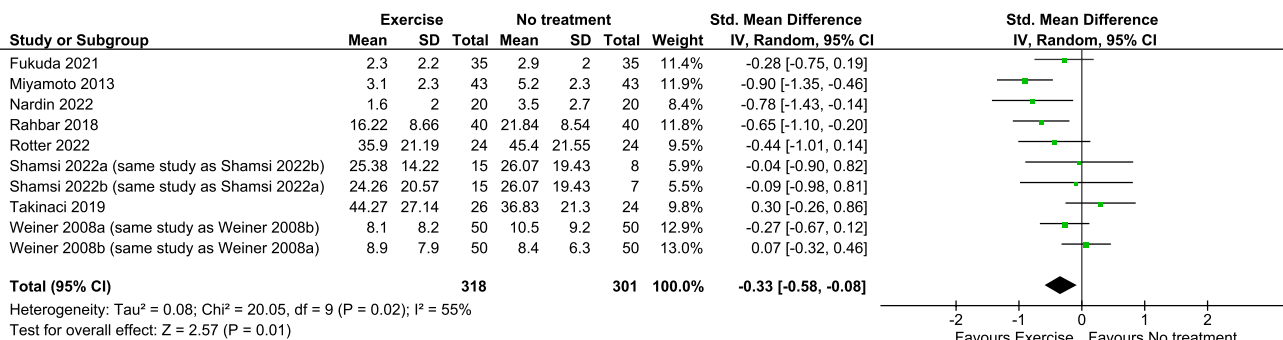Favours Exercise Favours No treatment

Fig. 2 Any exercise versus comparison interventions where the attributable effect of exercise could be isolated for pain in the immediate term (closest to 2 weeks)
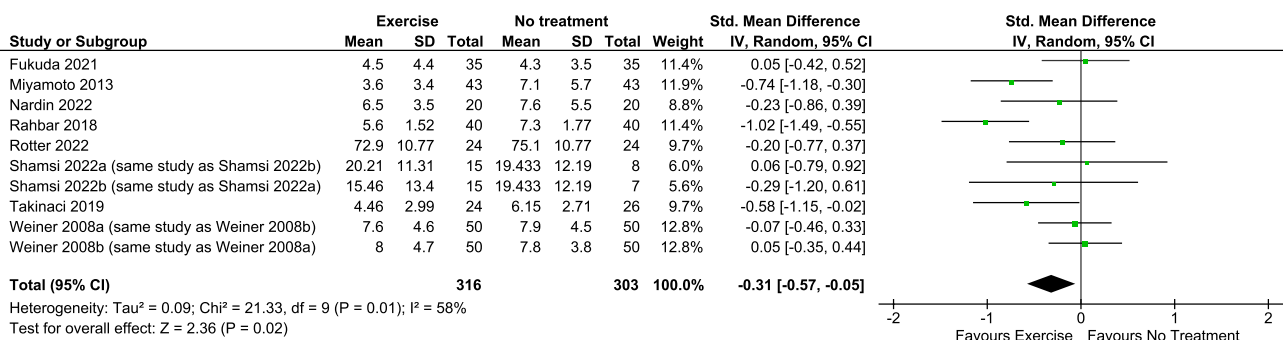
Study or Subgroup	Exercise Mean	SD	Total	No treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Fukuda 2021	4.5	4.4	35	4.3	3.5	35	11.4%	0.05 [-0.42, 0.52]
Miyamoto 2013	3.6	3.4	43	7.1	5.7	43	11.9%	-0.74 [-1.18, -0.30]
Nardin 2022	6.5	3.5	20	7.6	5.5	20	8.8%	-0.23 [-0.86, 0.39]
Rahbar 2018	5.6	1.52	40	7.3	1.77	40	11.4%	-1.02 [-1.49, -0.55]
Rotter 2022	72.9	10.77	24	75.1	10.77	24	9.7%	-0.20 [-0.77, 0.37]
Shamsi 2022a (same study as Shamsi 2022b)	20.21	11.31	15	19.433	12.19	8	6.0%	0.06 [-0.79, 0.92]
Shamsi 2022b (same study as Shamsi 2022a)	15.46	13.4	15	19.433	12.19	7	5.6%	-0.29 [-1.20, 0.61]
Takinaci 2019	4.46	2.99	24	6.15	2.71	26	9.7%	-0.58 [-1.15, -0.02]
Weiner 2008a (same study as Weiner 2008b)	7.6	4.6	50	7.9	4.5	50	12.8%	-0.07 [-0.46, 0.33]
Weiner 2008b (same study as Weiner 2008a)	8	4.7	50	7.8	3.8	50	12.8%	0.05 [-0.35, 0.44]
Total (95% CI)			316			303	100.0%	-0.31 [-0.57, -0.05]

Heterogeneity: Tau² = 0.09; Chi² = 21.33, df = 9 (P = 0.01); I² = 58%
Test for overall effect: Z = 2.36 (P = 0.02)

Favours Exercise Favours No Treatment

Fig. 3 Any exercise versus comparison interventions where the attributable effect of exercise could be isolated for pain in the immediate term (closest to 2 weeks)

lower values) (plot 9.1.2.2) [28, 31], or long term (1 RCT; between-group MD = 1.62, 95% CI − 0.06 to 3.31; *scale 0 to 24, 0 = no functional limitations*) (no plot, narrative synthesis) [30]. Exercise may make little or no difference to function in the intermediate term (4 RCTs; SMD = − 0.16, 95% CI − 0.39 to 0.07; *benefit indicated by lower values;* low certainty evidence) (plot 9.1.2.3) [23, 25, 31, 33].

In the immediate term, it is uncertain whether exercise makes little or no difference to ***HRQoL*** *(scale 0 to 100, 0 = poor quality of life)* for the ***PCS*** (2 RCTs; MD = − 2.31, 95% CI − 10.36 to 5.75; very low certainty evidence) (plot 9.1.3.1.1) [28, 33]. Exercise may make little or no difference for the ***MSC*** (2 RCTs; MD = − 1.11, 95% CI − 3.67 to 1.45; low certainty evidence) (plot 9.1.3.1.2) [28, 33]. It is uncertain whether exercise makes little or no difference to ***harms*** (2 RCTs; OR = 4.24, 95% CI 0.69 to 25.95; very low certainty evidence) (plot 9.1.4) [30, 33]. Harms were minor adverse events including back and knee pain.

Subgroup, Sensitivity and Supplementary Evidence Analyses

For the primary evidence synthesis, we did not conduct subgroup analysis for exercise dosage or intensity because there were too few RCTs (1–3) per comparison with little variation in dosage or intensity between RCTs. Additionally, we did not conduct sensitivity analyses removing the overall unclear ROB RCTs as most were given this rating (11/13, 85%).

In the supplementary evidence synthesis (see Online Resource 8), our findings aligned with our primary synthesis, except the certainty of evidence was lower (due to including RCTs rated as having high overall ROB). The supplementary evidence synthesis included 68 RCTs (13 identified from our primary synthesis [21–33], and 55 identified for the supplementary synthesis [37–91]. The 68 RCTs included a total of 4195 participants (ranging from 14 to 313 participants per RCT). The trials were conducted in high to upper-middle income economies: Australia (3 RCTs) [22, 72, 90], Brazil (7 RCTs) [23, 25, 26, 47, 48, 74, 77], Canada (2 RCTs) [64, 67], China (8 RCTs) [53, 54, 68–71, 83, 88], France (1 RCT) [78], Germany (3 RCTs) [28, 49, 50], Italy (1 RCT) [45], Japan (1 RCT) [24], Malaysia (1 RCT) [37], Netherlands (1 RCT) [29], South Korea (4 RCTs) [65, 66, 76, 86], Thailand (1 RCT) [59], Turkey (1 RCT) [31], United Kingdom (1 RCT) [32], and the United States (2 RCTs) [33, 46]; and low to lower-middle income economies: Egypt (2 RCTs) [37, 52], India (5 RCTs) [21, 51, 55, 79, 87], Iran (18 RCTs) [27, 29, 40, 42–44, 56–58, 62, 63, 72, 75, 81, 81, 83, 84, 89], Nigeria (2 RCTs) [59, 91], and Pakistan (4 RCTs) [39, 41, 61, 80]. The mean age of participants ranged from 20.4 to 74.3 years; nine RCTs with 524 participants total assessed older adults aged ≥ 60 years [24, 33, 45, 48, 54, 72, 76, 89, 90].

In the subgroup and/or sensitivity analyses conducted in both the primary and supplementary evidence syntheses, for all comparisons and outcomes, subgroup differences could not be explained and/or the differences between subgroups would likely not result in different recommendations for different subgroups. This was mostly due to the low or very low certainty evidence and the absence of or unimportant differences between the intervention and comparison groups (see Online Resources 7 and 8).

Discussion

The evidence regarding the benefits and harms of structured exercise programs for CPLBP in adults is based on 13 RCTs deemed as low or unclear ROB with a total of 1362 participants. Of these, two RCTs (n = 252) assessed adults aged ≥ 60 years. The eight exercise types assessed were aerobic exercise, core strengthening, muscle strengthening, mixed exercise, Pilates, stretching/flexibility/mobilizing exercise, yoga, and motor control exercise. Most of the RCTs (11, 85%) were rated as unclear overall ROB (concerns primarily with performance and detection bias). The certainty for the evidence related to individual exercise types was low or very low. Compared to no intervention, pain reduction was associated with aerobic exercise in the immediate and short terms, and Pilates in the immediate term, and motor control exercise vs. sham in the immediate and long terms. Improved function was associated with mixed exercise vs. usual care, and Pilates vs. no intervention in the immediate term. Temporary increased minor pain was associated with mixed exercise vs. no intervention, and yoga vs. usual care; no harms were reported with Pilates vs. no intervention. Little to no differences were found for other comparisons and outcomes.

When pooling all exercise types together based on the 13 RCTs, we found moderate certainty evidence indicating that in the immediate term, exercise (including aerobic, motor control, Pilates, yoga, core strengthening, and mixed exercise) improves pain in adults, and function in adults and older adults. Little or no difference was found between groups for the other outcomes (HRQoL, depression, self-efficacy, catastrophizing, fear avoidance, and performance-based physical functioning in older adults). Taken together, the findings from our primary synthesis, supplementary synthesis, and the work by Hayden et al. [2, 4] are consistent.

Our systematic review has several strengths. First, our international team had clinical and methodological expertise regarding LBP, systematic reviews, evidence syntheses, and answering important public health questions from the WHO. Second, our review process involved

conducting comprehensive literature searches without any language restrictions. Third, during the screening and ROB assessments, a core team member (with the most expertise and reliability in screening and ROB evaluations) was involved in each screening and ROB pair. Fourth, our ROB assessments did not rely on summary scores or the number of items at ROB. Instead, we created supplementary guidance forms based on the ROB 1 criteria [13, 14], which allowed reviewers to consider critical flaws in the studies [6]. Our use of these forms resulted in high agreement on ROB judgements. Fifth, we maintained transparency throughout the review process, providing detailed ROB assessments and footnotes for grading the certainty of the evidence (see Online Resources 2, 5, 8). These notes give readers a better understanding of our judgements and allow them to reach their own conclusions.

Our review has some limitations. One limitation is that we did not search the grey literature, which could introduce publication bias as studies published in peer-reviewed journals tend to report larger intervention effects than those in the grey literature [92]. We tried to mitigate this by searching for unpublished RCTs in the WHO ICTRP registry and contacting authors of unpublished RCTs. Moreover, unpublished studies are known to represent a small proportion of studies and rarely impact results and conclusions [93]. However, it may be important to include such studies in limited scenarios or where there are potential conflicts of interest in published research [93].

We identified several key gaps in the evidence across different exercise comparisons: 1) lack of studies examining the effects of exercise on anxiety symptoms and social participation (including work); 2) inability to assess whether the benefits or harms of exercise interventions vary by gender/sex or race/ethnicity; 3) insufficient studies to evaluate the impact of leg pain/symptoms on exercise benefits or harms, as well as differences in higher versus lower income countries; 4) inability to examine the influence of intervention-level characteristics, such as exercise specificity, tailored approaches, supervision level, and group versus individual delivery, on benefits and harms; 5) limited evidence on the benefits or harms of specific exercise types in older adults, including aerobic exercise, core strengthening, muscle strength training, Pilates, stretching, flexibility or mobilizing exercises, yoga, and motor control exercises; 6) few studies assessing the impact of exercise on quality of life and psychological outcomes (depression, fear avoidance, catastrophizing, self-efficacy, anxiety), with comparatively less evidence available for older individuals; 7) limited understanding of the effects of exercise in vulnerable populations, such as older adults and those in low-income settings, who are more likely to experience persistent disability from low back pain.

Additionally, exercise's effects are modest, suggesting a need for multifaceted interventions.

Conclusion

When assessing individual exercise types, based on low or very low certainty evidence, pain reduction was associated with aerobic exercise, Pilates and motor control exercise; improved function was associated with mixed exercise and Pilates. A temporary increase in minor pain was associated with mixed exercise and yoga. Little to no difference was found for other comparisons and outcomes. When pooling exercise types, based on moderate certainty evidence, exercise was shown to be beneficial in improving pain and function in adults and older adults. Exercise prescription should be considered based on patient preferences, availability of exercise type, costs, and other contextual factors. Harms should be further investigated systematically.

Data Availability The datasets generated during and/or analysed during the primary analysis of the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of Interests All team members provided DOI forms to WHO for evaluation at inception. CC (Carol Cancelliere), LV, DS, HY, GC, JJW, HMS report funding from the Canadian Chiropractic Guideline Initiative. JAH and RO report funding from the Canadian Institutes of Health Research to support the 'exercise treatment for chronic low back pain' Cochrane review. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. CAH reports grants to the University of Zurich from the Foundation for the Education of Chiropractors in Switzerland, the Swiss National Science Foundation (SNSF), and the European Centre for Chiropractic Research Excellence (ECCRE) outside the submitted work. SM reports grants from Canadian Chiropractic Association, Canadian Chiropractic Research Foundation, and Canadian Institute of Health Research outside of submitted work. AB reports grants from the Canadian Chiropractic Association, Canadian Chiropractic Research Foundation, and Health Canada outside the submitted work. JJW reports grants from the Canadian Institutes of Health Research (CIHR) and Canadian Chiropractic Research Foundation (paid to university), and travel reimbursement for research meetings from the Chiropractic Academy of Research Leadership outside the submitted work.

Ethics Ethical approval was not required for this systematic review of previously published studies.

References

1. Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, et al. The role of activity in the therapeutic management of back pain Report of the International Paris Task Force on Back Pain. Spine (Phila Pa 1976. 2000;25(4 Suppl):1S-33S.

2. Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. Cochrane Database Syst Rev. 2021;9:CD009790.

3. Powell KE, Paluch AE, Blair SN. Physical activity for health: What kind? How much? How intense? On top of what? Annu Rev Public Health. 2011;32:349–65.

4. Hayden JA, Ellis J, Ogilvie R, Stewart SA, Bagg MK, Stanojevic S, et al. Some types of exercise are more effective than others in people with chronic low back pain: a network meta-analysis. J Physiother. 2021;67(4):252–62.

5. Hayden JA, Group BEC-CRW. Commentary: collaborative systematic review may produce and share high-quality, comparative evidence more efficiently. J Clin Epidemiol. 2022;152:288–94.

6. Cancelliere C, Verville L, Southerst D, Yu H, Hayden JA, Ogilvie R, Bussières A, Gross DP, Pereira P, Mior S, Tricco AC, Cedraschi C, Brunton G, Nordin M, Wong JJ, Shearer HM, Connell G, Hincapié CA. Systematic review procedures for the World Health Organization (WHO) evidence syntheses on benefits and harms of structured and standardized education/advice, structured exercise programs, transcutaneous electrical nerve stimulation (TENS), and needling therapies for the management of chronic low back pain in adults. J Occup Rehabil. 2023. https://doi.org/10.1007/s10926-023-10156-w.

7. Hayden JA, Ogilvie R, Kashif S, Singh S, Boulos L, Stewart S, et al. Exercise treatments for chronic low back pain: a network meta-analysis (protocol). Cochrane Database Syst Rev. Submitted revision, September 2022.

8. Herbert RD. How to estimate treatment effects from reports of clinical trials. I: Continuous outcomes. Aust J Physiother. 2000;46(3):229–35.

9. Herbert RD. How to estimate treatment effects from reports of clinical trials. II: Dichotomous outcomes. Aust J Physiother. 2000;46(4):309–13.

10. Rubinstein SM, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. Cochrane Database Syst Rev. 2012;2012(9):CD008880.

11. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain. 2008;9(2):105–21.

12. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. New Jersey (NJ): Lawrence Erlbaum Associates; 1988.

13. Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

14. Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 updated method guideline for systematic reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015;40(21):1660–73.

15. DistillerSR. Version 2023.2.1. 2023.2.1 ed. Ottawa, Canada: DistillerSR; 2023.

16. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ. 2020;368:l6890.

17. Schünemann HJ, Higgins JPT, Vist GE, Glasziou P, Akl EA, Skoetz N, et al. Chapter 14: Completing 'Summary of findings' tables and grading the certainty of the evidence. 2022 [cited January 5, 2022]. In: Cochrane Handbook for Systematic Reviews of Interventions version 63 (updated February 2022) [Internet]. Cochrane, [cited January 5, 2022]. Available from: www.training. cochrane.org/handbook.

18. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria2022.

19. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softwe. 2010;36(3):1–48.

20. GRADEpro. GDT: GRADEpro Guideline Development Tool (Software). McMaster University and Evidence Prime; 2021.

21. Chhabra HS, Sharma S, Verma S. Smartphone app in self-management of chronic low back pain: a randomized controlled trial. Eur Spine J. 2018;27(11):2862–74.

22. Costa LO, Maher CG, Latimer J, Hodges PW, Herbert RD, Refshauge KM, et al. Motor control exercise for chronic low back pain: a randomized placebo-controlled trial. Phys Ther. 2009;89(12):1275–86.

23. Fukuda TY, Aquino LM, Pereira P, Ayres I, Feio AF, de Jesus FLA, et al. Does adding hip strengthening exercises to manual therapy and segmental stabilization improve outcomes in patients with nonspecific low back pain? A randomized controlled trial. Braz J Phys Ther. 2021;25(6):900–7.

24. Jinnouchi H, Matsudaira K, Kitamura A, Kakihana H, Oka H, Hayama-Terada M, et al. Effects of brief self-exercise education on the management of chronic low back pain: a community-based, randomized, parallel-group pragmatic trial. Mod Rheumatol. 2020;31(4):890–8.

25. Miyamoto GC, Costa LO, Galvanin T, Cabral CM. Efficacy of the addition of modified Pilates exercises to a minimal intervention in patients with chronic low back pain: a randomized controlled trial. Phys Ther. 2013;93(3):310–20.

26. Nardin DMK, Stocco MR, Aguiar AF, Machado FA, de Oliveira RG, Andraus RAC. Effects of photobiomodulation and deep water running in patients with chronic non-specific low back pain: a randomized controlled trial. Lasers Med Sci. 2022;37(4):2135–44.

27. Rahbar M, Salekzamani Y, Jahanjou F, Eslamian F, Niroumand A, Dolatkhah N. Effect of hippotherapy simulator on pain, disability and range of motion of the spinal column in subjects with mechanical low back pain: a randomized single-blind clinical trial. J Back Musculoskelet Rehabil. 2018;31(6):1183–92.

28. Rotter G, Ortiz M, Binting S, Tomzik J, Reese F, Roll S, et al. Mindful walking in patients with chronic low back pain: a randomized controlled trial. J Integr Complement Med. 2022;28(6):474–83.

29. Shamsi M, Ahmadi A, Mirzaei M, Jaberzadeh S. Effects of static stretching and strengthening exercises on flexion relaxation ratio in patients with LBP: a randomized clinical trial. J Bodyw Mov Ther. 2022;30:196–202.

30. Smeets R, Vlaeyen JWS, Hidding A, Kester ADM, van der Heijden G, Knottnerus AJ. Chronic low back pain: physical training, graded activity with problem solving training, or both? The one-year post-treatment results of a randomized controlled trial. Pain. 2008;134(3):263–76.

31. Takinaci Z. Clinical efficacy of spa therapy (balneotherapy) for chronic low back pain: a randomized single-blind trial. Eur J Integr Med. 2019;29:100928.

32. Tilbrook HE, Cox H, Hewitt CE, Kang'ombe AR, Chuang LH, Jayakody S, et al. Yoga for chronic low back pain: a randomized trial. Ann Intern Med. 2011;155(9):569–78.

33. Weiner DK, Perera S, Rudy TE, Glick RM, Shenoy S, Delitto A. Efficacy of percutaneous electrical nerve stimulation and therapeutic exercise for older adults with chronic low back pain: a randomized controlled trial. Pain. 2008;140(2):344–57.

34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81.

35. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform. 2019;95:103208.

36. World Bank Country and Lending Groups. [cited 2022 May 1, 2022]. Available from: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups.

37. Abadi FH, Sankaravel M, Zainuddin FF, Elumalai G, Razli AI. The effect of aquatic exercise program on low-back pain disability in obese women. J Exerc Rehabil. 2019;15(6):855–60.

38. Abdel-Aziem AA, Abdelraouf OR, El-Basatiny HMY, Draz AH. The effects of stabilization exercises combined with pelvic floor exercise in women with nonspecific low back pain: a randomized clinical study. J Chiropr Med. 2021;20(4):229–38.

39. Afzal MW, Ahmad A, Mohseni Bandpei MA, Gilani SA, Hanif A, Waqas MS. Effects of virtual reality exercises and routine physical therapy on pain intensity and functional disability in patients with chronic low back pain. J Pak Med Assoc. 2022;72(3):413–7.

40. Ahmadizadeh Z, Ehsani F, Samaei SA, Mirmohammadkhani M. The effect of stabilization exercises along with self-care training on transverse abdominal activity, pain, and disability in mothers with low back pain having children with cerebral palsy: a randomized controlled trial. Am J Phys Med Rehabil. 2020;99(2):156–60.

41. Ain SQT, Rehman SSU, Maryam M, Kiani SK. Effects of sustained natural apophyseal glides with and without thoracic posturecorrection techniques on mechanical back pain: a randomized control trial. J Pak Med Assoc. 2019;69(11):1584–7.

42. Alikhajeh Y, Barabadi E, Rahimi GRM. A comparison of 6 weeks of aquatic exercise and Kinesio taping in patients with chronic nonspecific low back pain. J Sport Rehabil. 2021;30:37–42.

43. Alvani E, Shirvani H, Shamsoddini A. Neuromuscular exercises on pain intensity, functional disability, proprioception, and balance of military personnel with chronic low back pain. J Can Chiropr Assoc. 2021;65(2):193–206.

44. Ansari S, Elmieh A, Alipour A. The effect of aquatic exercise on functional disability, flexibility and function of trunk muscles in postmenopausal women with chronic non-specific low back pain: randomized controlled trial. Sci Sports. 2021;36(3):e103–10.

45. Barni L, Calabretta L, Lepori L, Pasquetti P, Gulisano M, Freddolini M. Does an aerobic exercise improve outcomes in older sedentary nonspecific low back pain subjects? A randomized controlled study. Top Geriatr Rehabil. 2018;34(2):88–94.

46. Bruehl S, Burns JW, Koltyn K, Gupta R, Buvanendran A, Edwards D, et al. Are endogenous opioid mechanisms involved in the effects of aerobic exercise training on chronic low back pain? A randomized controlled trial. Pain. 2020;161(12):2887–97.

47. Dimer da Luz R, da Silva SM, Steffen Evaldt A, da Silva ML, Boff Daitx R, Dohnert MB. Neuromuscular electrical stimulation associated with core stability exercises in nonspecific postural low back pain: a randomized clinical trial. Muscles Ligaments Tendons J. 2019;9(3):446–556.

48. da Silva TMJC, da Silva NN, de Souza Rocha SH, de Oliveira DM, Monte-Silva KK, de Silva TA, et al. Back school program for back pain: education or physical exercise? ConScientiae Saúde. 2014;13(4):506–15.

49. Dalichau S, Scheele K. Status of aquatic functional training in the therapy of chronic back pain. Phys Med Rehab Kuror. 2003;13:35–41.

50. Dalichau S, Stein B, Schafer K, Buhlmann JJ, Menken P. Quantification of spinal configuration and postural capacity by ultrasound topometry for evaluation of different muscle strengthening programs in the therapy of back pain. Z Orthop. 2005;143:79–85.

51. Dineshkumar SK, Dibyendunarayan B, Ramalingam AT. Effect of abdominal drawing-in maneuver along with resisted ankle dorsiflexion to activate transverse abdominis muscle in chronic nonspecific low back pain. Rom J Phys Ther. 2015;21(36):40–7.

52. Fouda KZ, Dewir IM, Abdelsalam MS. Effects of proprioceptive neuromuscular facilitation techniques in treating chronic nonspecific low back pain patients. Physiother Quart. 2021;29(2):32–7.

53. Gao F, Yuan S, Liu N, Zheng S. Effect observation of Baduanjin training on chronic nonspecific low back pain. Rehabil Med. 2018;28(4):13–7.

54. Ge L, Huang H, Yu Q, Li Y, Li X, Li Z, et al. Effects of core stability training on older women with low back pain: a randomized controlled trial. Eur Rev Aging Phys Act. 2022;19(1):10.

55. Gupta P, Mohanty PP, Pattnaik M. The effectiveness of aerobic exercise program for improving functional performance and quality of life in chronic low back pain. Indian J Physiother Occup Ther. 2019;13(2):155–60.

56. Hatefi M, Babakhani F, Ashrafizadeh M. The effect of static stretching exercises on hip range of motion, pain, and disability in patients with non-specific low back pain. J Exp Orthop. 2021;8(1):55.

57. Heidari RS, Sahebozamani M, Karimi AF. Comparison of the effects of 8 weeks of core stability exercise on ball and sling exercise on the quality of life and pain in the female with non-specific chronic low back pain (NSLBP). J Zanjan University of Medical Sciences & Health Services. 2018.

58. Hemmati S, Rajabi R, Karimi N, Jahandideh A. Effects of consecutive supervised core stability training on pain and disability in women with nonspecific chronic low back pain. Koomesh. 2011;12(3):244–52.

59. Idowu OA, Adeniyi AF. Efficacy of graded activity with and without daily-monitored-walking on pain and back endurance among patients with concomitant low-back pain and type-2 diabetes: A randomized trial. Ethiop J Health Sci. 2020;30(2):233–42.

60. Joseph LH, Hancharoenkul B, Sitilertpisan P, Pirunsan U, Paungmali A. Effects of massage as a combination therapy with lumbopelvic stability exercises as compared to standard massage therapy in low back pain: A randomized cross-over study. Int J Ther Massage Bodywork. 2018;11(4):16–22.

61. Kanwal S, Yaqoob I, Shakil-Ur-Rehman S, Ghous M, Ghazal J, Namroz N. Effects of core muscle stability on low back pain and quality of life in post- menopausal women: a comparative study. J Pak Med Assoc. 2021;71(1(A)):37–40.

62. Karimi N, Talimkhani A, Mosallanezhad Z, Arab AM, Keshavarz R. The effects of consecutive supervised functional lumbar stabilizing exercises on the postural balance and functional disability in low back pain. Iran Rehabil J. 2014;12(20):21–7.

63. Karimzadeh F, Ghasemi GA, Letafatkar A. The effect of 8 weeks core stabilization exercises on pain and functional disability induced by low back pain in the mothers of children with cerebral palsy. Sci J Kurdistan Univ Med Sci. 2016;21(44):5931–94.

64. Kell RT, Risi AD, Barden JM. The response of persons with chronic nonspecific low back pain to three different volumes of periodized musculoskeletal rehabilitation. J Strength Cond Res. 2011;25(4):1052–64.

65. Kim S, Jee Y. Effects of 3D moving platform exercise on physiological parameters and pain in patients with chronic low back pain. Medicina (Kaunas). 2020;56(7):1–15.

66. Kim B, Yim J. Core stability and hip exercises improve physical function and activity in patients with non-specific low back pain: a randomized controlled trial. Tohoku J Exp Med. 2020;251(3):193–206.

67. Lang AE, Hendrick PA, Clay L, Mondal P, Trask CM, Bath B, et al. A randomized controlled trial investigating effects of an individualized pedometer driven walking program on chronic low back pain. BMC Musculoskelet Disord. 2021;22(1):206.

68. Li L, Feng Z. Clinical observation of treating chronic nonspecific low back pain with Baduanjin combining sling exercise therapy. Rheumatism and Arthritis. 2015;4(4):16–7.

69. Li Z, Yu Q, Luo H, Liang W, Li X, Ge L, et al. The effect of virtual reality training on anticipatory postural adjustments in patients with chronic nonspecific low back pain: a preliminary study. Neural Plast. 2021;2021:9975862.

70. Liu J, Zhao W, Yuan Y. Effects of Tai Chi on the event-related potential of patients with chronic non-specific low back pain. Chin J Sports Med. 2018;37(10):826–32.

71. Liu J, Yeung A, Xiao T, Tian X, Kong Z, Zou L, et al. Chen-style Tai Chi for individuals (aged 50 years old or above) with chronic non-specific low back pain: a randomized controlled trial. Int J Environ Res Public Health. 2019;16(3):517.

72. Madadi-Shad M, Jafarnezhadgero AA, Sheikhalizade H, Dionisio VC. Effect of a corrective exercise program on gait kinetics and muscle activities in older adults with both low back pain and pronated feet: a double-blind, randomized controlled trial. Gait Posture. 2020;76:339–45.

73. McIlveen B, Robertson VJ. A randomised controlled study of the outcome of hydrotherapy for subjects with low back or back and leg pain. Physiotherapy. 1998;84(1):17–26.

74. Mendes PRF, Gomes SRA, Costa LDO, Liguori A, Bulhoes LCC, Brasileiro JS. Core stabilisation exercises reduce chronic low back pain in Air Force fighter pilots: a randomised controlled trial. BMJ Mil Health. 2022.

75. Noormohammadpour P, Kordi M, Mansournia MA, Akbari-Fakhrabadi M, Kordi R. The role of a multi-step core stability exercise program in the treatment of nurses with chronic low back pain: A single-blinded randomized controlled trial. Asian Spine J. 2018;12(3):490–502.

76. Park S, Park S, Min S, Kim C-J, Jee Y-S. A randomized controlled trial investigating the effects of equine simulator riding on low back pain, morphological changes, and trunk musculature in elderly women. Medicina. 2020;56(610):1–14.

77. Prado ERA, Meireles SM, Carvalho ACA, Mazoca MF, Motta Neto AM, Barboza Da Silva R, et al. Influence of isostretching on patients with chronic low back pain. A randomized controlled trial. Physiother Theory Pract. 2021;37(2):287–94.

78. Raoul T, Malfériot J, Barizien N, Lhuissier FJ. Effects of a spinal extensor muscle strengthening program in chronic low back pain triathletes. Randomized controlled study in 67 athletes. J Traumatologie du Sport. 2019;36(3):183–93.

79. Rathi M. Effect of pelvic floor muscle strengthening exercises in chronic low back pain. Indian J Physiother Occup Ther. 2013;7(1):121–5.

80. Raza S, Awan WA, Ghauri MW, Mahmood T, Abbas S. Effectiveness of spinal stabilization exercises with and without stretching of Latissimus dorsi Muscle in chronic mechanical low back pain. Rawal Med J. 2020;45(4):857–62.

81. Sedaghati P, Ardjmand A, Sedaghati N. Comparison of the effects of different training approaches on dynamic balance and pain intensity in the patients with chronic back pain. Sci J Kurdistan Univ Medical Sci. 2017;90:6–13.

82. Shariat A, Alizadeh R, Moradi V, Afsharnia E, Hakakzadeh A, Ansari NN, et al. The impact of modified exercise and relaxation therapy on chronic lower back pain in office workers: a randomized clinical trial. J Exerc Rehabil. 2019;15(5):703–8.

83. Shu J, Li D, Tao W, Chen S. Observation on the curative effect of massage manipulation combined with core strength training in patients with chronic nonspecific low back pain. Evid Based Complement Alternat Med. 2021;2021:7534577.

84. Sokhangoi Y, Noorabadi MH, Noorabadi MS. The effect of Pilates exercises on chronic non-specific low back pain and stable range of motion in women 40 to 60 years old. Sci Res J Zanjan Univ Med Sci. 2017;25(108):20–30.

85. Srivastav N, Joshi S, Kushwah SS. Comparison between effectiveness of lumbar stabilization exercises and conventional physical therapy in the management of mechanical low back pain Indian. J Physiother Occup Ther. 2018;12(4):28–33.

86. Suh JH, Kim H, Jung GP, Ko JY, Ryu JS. The effect of lumbar stabilization and walking exercises on chronic low back pain: a randomized controlled trial. Medicine (Baltimore). 2019;98(26):e16173.

87. Wattamwar RB, Nadkarni K. Effect of conventional occupational therapy and yoga in chronic low back pain. Indian J Occup Ther. 2012;45(3):13–20.

88. Xu C, Fu Z, Wang X. Effect of transversus abdominis muscle training on pressure-pain threshold in patients with chronic low back pain. BMC Sports Sci Med Rehabil. 2021;13(1):35.

89. Yalfani A, Abedi M, Raeisi Z. Effects of an 8-week virtual reality training program on pain, fall risk, and quality of life in elderly women with chronic low back pain: Double-blind randomized clinical trial. Games Health J. 2022;11(2):85–92.

90. Zadro JR, Shirley D, Simic M, Mousavi SJ, Ceprnja D, Maka K, et al. Video-game-based exercises for older people with chronic low back pain: a randomized controlled trial (GAMEBACK). Phys Ther. 2019;99(1):14–27.

91. Zakari UU, Bello B, Sokumbi GO, Yakasai AM, Danazumi MS. Comparison of the effects of positional release therapy and lumbar stabilization exercises in the management of chronic mechanical low back pain - Randomized controlled trial. Crit Rev Phys Rehabil Med. 2019;31(4):321–32.

92. Hopewell S, McDonald S, Clark MJ, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. Cochrane Database Syst Rev. 2007;2007(2):MR000010.

93. Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. BMC Med Res Methodol. 2017;17(1):64.

Authors and Affiliations

Leslie Verville[1] · Rachel Ogilvie[2] · Cesar A. Hincapié[3,4,5] · Danielle Southerst[1] · Hainan Yu[1] · André Bussières[6,7] · Douglas P. Gross[8] · Paulo Pereira[9] · Silvano Mior[1,10] · Andrea C. Tricco[11,12,13] · Christine Cedraschi[14,15] · Ginny Brunton[1,16,17] · Margareta Nordin[18] · Gaelan Connell[1] · Jessica J. Wong[1] · Heather M. Shearer[1,10,19] · Joyce G. B. Lee[10] · Dan Wang[1] · Jill A. Hayden[2] · Carol Cancelliere[1]

✉ Cesar A. Hincapié
cesar.hincapie@uzh.ch

✉ Carol Cancelliere
carolina.cancelliere@ontariotechu.ca

1 Institute for Disability and Rehabilitation Research and Faculty of Health Sciences, Ontario Tech University, Oshawa, Canada

2 Department of Community Health and Epidemiology, Dalhousie University, Halifax, Canada

3 EBPI-UWZH Musculoskeletal Epidemiology Research Group, University of Zurich and Balgrist University Hospital, Zurich, Switzerland

4 Epidemiology, Biostatistics and Prevention Institute (EBPI), University of Zurich, Zurich, Switzerland

5 University Spine Centre Zurich (UWZH), Balgrist University Hospital and University of Zurich, Zurich, Switzerland

6 Département Chiropratique, Université du Québec à Trois-Rivières, Trois-Rivières (Québec), Canada

7 School of Physical and Occupational Therapy, Faculty of Medicine and Health Sciences, McGill University, Montreal, Canada

8 Department of Physical Therapy, University of Alberta, Edmonton, Canada

9 Department of Neurosurgery, Centro Hospitalar Universitário São João, Faculty of Medicine, University of Porto, Porto, Portugal

10 Department of Research and Innovation, Canadian Memorial Chiropractic College, Toronto, Canada

11 Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, Toronto, Canada

12 Epidemiology Division and Institute for Health Policy, Management, and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

13 Queen's Collaboration for Health Care Quality Joanna Briggs Institute Centre of Excellence, Queen's University, Kingston, Canada

14 Division of General Medical Rehabilitation, Geneva University and University Hospitals, Geneva, Switzerland

15 Division of Clinical Pharmacology and Toxicology, Multidisciplinary Pain Centre, Geneva University Hospitals, Geneva, Switzerland

16 EPPI-Centre, UCL Institute of Education, University College London, London, England, UK

17 McMaster Midwifery Research Centre, McMaster University, Hamilton, Canada

18 Department of Orthopedic Surgery and Environmental Medicine, NYU Grossman School of Medicine, New York University, New York, United States

19 Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Canada