Cochrane Library

Cochrane Database of Systematic Reviews

Flexibility exercise training for adults with fibromyalgia (Review)

Kim SY, Busch AJ, Overend TJ, Schachter CL, van der Spuy I, Boden C, Góes SM, Foulds HJA, Bidonde J

Kim SY, Busch AJ, Overend TJ, Schachter CL, van der Spuy I, Boden C, Góes SM, Foulds HJA, Bidonde J.
Flexibility exercise training for adults with fibromyalgia.
Cochrane Database of Systematic Reviews 2019, Issue 9. Art. No.: CD013419.
DOI: 10.1002/14651858.CD013419.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

HEADER 1
ABSTRACT 1
PLAIN LANGUAGE SUMMARY 2
SUMMARY OF FINDINGS 4
BACKGROUND 7
OBJECTIVES 8
METHODS 8
RESULTS 13
Figure 1. 14
Figure 2. 17
Figure 3. 18
Figure 4. 21
DISCUSSION 24
AUTHORS' CONCLUSIONS 27
ACKNOWLEDGEMENTS 28
REFERENCES 29
CHARACTERISTICS OF STUDIES 36
DATA AND ANALYSES 57
Analysis 1.1. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention). 58
Analysis 1.2. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention). 58
Analysis 1.3. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention). 58
Analysis 1.4. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 4 Stiffness, 0-100, lower is best (end of intervention). 58
Analysis 1.5. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 5 Physical function, 0-100, lower is best (end of intervention). 59
Analysis 1.6. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 6 Depression, 0-100, lower is best (end of intervention). 59
Analysis 1.7. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 7 Tenderness 0-18, lower is best (end of intervention). 59
Analysis 1.8. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 8 Withdrawals. 59
Analysis 1.9. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 9 Long-term effects. 60
Analysis 2.1. Comparison 2 Flexibility versus control (end of intervention), Outcome 1 Pain, Intensity, 0-100, lower is best (end of intervention). 60
Analysis 2.2. Comparison 2 Flexibility versus control (end of intervention), Outcome 2 Physical function, 0-100, lower is best (end of intervention). 61
Analysis 2.3. Comparison 2 Flexibility versus control (end of intervention), Outcome 3 Withdrawals. 61
Analysis 3.1. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention). 62
Analysis 3.2. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention). 62
Analysis 3.3. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention). 62
Analysis 3.4. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 4 Physical function, 0-100, lower is best (end of intervention). 62
Analysis 3.5. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 5 Depression, 0-63, lower is best (end of intervention). 63
Analysis 3.6. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 6 Tenderness, 0-18, lower is best (end of intervention). 63
Analysis 3.7. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 7 > 30% improvement of pain (end of intervention). 63
Analysis 3.8. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 8 Withdrawals. 63

Analysis 4.1. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention). 65
Analysis 4.2. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention). 66
Analysis 4.3. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention). 66
Analysis 4.4. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 4 Stiffness, 0-100, lower is best (end of intervention). 66
Analysis 4.5. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 5 Physical function, 0-100, lower is best (end of intervention). 67
Analysis 4.6. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 6 Depression, 0-63, lower is best (end of intervention). 67
Analysis 4.7. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 7 Tenderness, 0-18, lower is best (end of intervention). 67
Analysis 4.8. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 8 Withdrawals. 68
Analysis 4.9. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 9 Long-term effects: flexibility vs other comparators. 68
ADDITIONAL TABLES 68
APPENDICES 82
HISTORY 90
CONTRIBUTIONS OF AUTHORS 90
DECLARATIONS OF INTEREST 91
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 91
NOTES 91
INDEX TERMS 92

[Intervention Review]

Flexibility exercise training for adults with fibromyalgia

Soo Y Kim[1], Angela J Busch[1], Tom J Overend[2], Candice L Schachter[1], Ina van der Spuy[3], Catherine Boden[4], Suelen M Góes[1], Heather JA Foulds[5], Julia Bidonde[6]

[1]School of Rehabilitation Science, University of Saskatchewan, Saskatoon, Canada. [2]School of Physical Therapy, University of Western Ontario, London, Canada. [3]School of Physical Therapy, University of Saskatchewan, Saskatoon, Canada. [4]Leslie and Irene Dube Health Sciences Library, University Library, University of Saskatchewan, Saskatoon, Canada. [5]College of Kinesiology, University of Saskatchewan, Saskatoon, Canada. [6]Norwegian Institute of Public Health, Oslo, Norway

Contact address: Soo Y Kim, School of Rehabilitation Science, University of Saskatchewan, Health Sciences Building, 104 Clinic Place, Room 3410, Saskatoon, SK, S7N 2Z4, Canada. soo.kim@usask.ca.

Editorial group: Cochrane Musculoskeletal Group
Publication status and date: New, published in Issue 9, 2019.

Citation: Kim SY, Busch AJ, Overend TJ, Schachter CL, van der Spuy I, Boden C, Góes SM, Foulds HJA, Bidonde J. Flexibility exercise training for adults with fibromyalgia. *Cochrane Database of Systematic Reviews* 2019, Issue 9. Art. No.: CD013419. DOI: 10.1002/14651858.CD013419.

ABSTRACT

Background

Exercise training is commonly recommended for adults with fibromyalgia. We defined flexibility exercise training programs as those involving movements of a joint or a series of joints, through complete range of motion, thus targeting major muscle-tendon units. This review is one of a series of reviews updating the first review published in 2002.

Objectives

To evaluate the benefits and harms of flexibility exercise training in adults with fibromyalgia.

Search methods

We searched the Cochrane Library, MEDLINE, Embase, CINAHL (Cumulative Index to Nursing and Allied Health Literature), PEDro (Physiotherapy Evidence Database), Thesis and Dissertation Abstracts, AMED (Allied and Complementary Medicine Database), the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP), and ClinicalTrials.gov up to December 2017, unrestricted by language, and we reviewed the reference lists of retrieved trials to identify potentially relevant trials.

Selection criteria

We included randomized trials (RCTs) including adults diagnosed with fibromyalgia based on published criteria. Major outcomes were health-related quality of life (HRQoL), pain intensity, stiffness, fatigue, physical function, trial withdrawals, and adverse events.

Data collection and analysis

Two review authors independently selected articles for inclusion, extracted data, performed 'Risk of bias' assessments, and assessed the certainty of the body of evidence for major outcomes using the GRADE approach. All discrepancies were rechecked, and consensus was achieved by discussion.

Main results

We included 12 RCTs (743 people). Among these RCTs, flexibility exercise training was compared to an untreated control group, land-based aerobic training, resistance training, or other interventions (i.e. Tai Chi, Pilates, aquatic biodanza, friction massage, medications). Studies were at risk of selection, performance, and detection bias (due to lack of adequate randomization and allocation concealment, lack of

participant or personnel blinding, and lack of blinding for self-reported outcomes). With the exception of withdrawals and adverse events, major outcomes were self-reported and were expressed on a 0-to-100 scale (lower values are best, negative mean differences (MDs) indicate improvement). We prioritized the findings of flexibility exercise training compared to land-based aerobic training and present them fully here.

Very low-certainty evidence showed that compared with land-based aerobic training, flexibility exercise training (five trials with 266 participants) provides no clinically important benefits with regard to HRQoL, pain intensity, fatigue, stiffness, and physical function. Low-certainty evidence showed no difference between these groups for withdrawals at completion of the intervention (8 to 20 weeks).

Mean HRQoL assessed on the Fibromyalgia Impact Questionnaire (FIQ) Total scale (0 to 100, higher scores indicating worse HRQoL) was 46 mm and 42 mm in the flexibility and aerobic groups, respectively (2 studies, 193 participants); absolute change was 4% worse (6% better to 14% worse), and relative change was 7.5% worse (10.5% better to 25.5% worse) in the flexibility group. Mean pain was 57 mm and 52 mm in the flexibility and aerobic groups, respectively (5 studies, 266 participants); absolute change was 5% worse (1% better to 11% worse), and relative change was 6.7% worse (2% better to 15.4% worse). Mean fatigue was 67 mm and 71 mm in the aerobic and flexibility groups, respectively (2 studies, 75 participants); absolute change was 4% better (13% better to 5% worse), and relative change was 6% better (19.4% better to 7.4% worse). Mean physical function was 23 points and 17 points in the flexibility and aerobic groups, respectively (1 study, 60 participants); absolute change was 6% worse (4% better to 16% worse), and relative change was 14% worse (9.1% better to 37.1% worse). We found very low-certainty evidence of an effect for stiffness. Mean stiffness was 49 mm to 79 mm in the flexibility and aerobic groups, respectively (1 study, 15 participants); absolute change was 30% better (8% better to 51% better), and relative change was 39% better (10% better to 68% better). We found no evidence of an effect in all-cause withdrawal between the flexibility and aerobic groups (5 studies, 301 participants). Absolute change was 1% fewer withdrawals in the flexibility group (8% fewer to 21% more), and relative change in the flexibility group compared to the aerobic training intervention group was 3% fewer (39% fewer to 55% more). It is uncertain whether flexibility leads to long-term effects (36 weeks after a 12-week intervention), as the evidence was of low certainty and was derived from a single trial.

Very low-certainty evidence indicates uncertainty in the risk of adverse events for flexibility exercise training. One adverse effect was described among the 132 participants allocated to flexibility training. One participant had tendinitis of the Achilles tendon (McCain 1988), but it is unclear if the tendinitis was a pre-existing condition.

Authors' conclusions

When compared with aerobic training, it is uncertain whether flexibility improves outcomes such as HRQoL, pain intensity, fatigue, stiffness, and physical function, as the certainty of the evidence is very low. Flexibility exercise training may lead to little or no difference for all-cause withdrawals. It is also uncertain whether flexibility exercise training has long-term effects due to the very low certainty of the evidence. We downgraded the evidence owing to the small number of trials and participants across trials, as well as due to issues related to unclear and high risk of bias (selection, performance, and detection biases). While flexibility exercise training appears to be well tolerated (similar withdrawal rates across groups), evidence on adverse events was scarce, therefore its safety is uncertain.

PLAIN LANGUAGE SUMMARY

Flexibility exercise training for adults with fibromyalgia

This review summarizes the effects of flexibility exercise for adults with fibromyalgia.

What problems do fibromyalgia cause?

People with fibromyalgia have persistent, widespread body pain. They may also have fatigue, anxiety, depression, and sleep difficulties.

What is flexibility exercise training?

Flexibility exercise training is a type of exercise that focuses on improving or maintaining the amount of motion available in muscles and joint structures by holding or stretching the body in specific positions.

Study characteristics

We searched the literature up to December 2017 and found 12 studies (743 individuals) that met our inclusion criteria. Flexibility interventions were compared with control (treatment as usual), aerobic training interventions (e.g. treadmill walking), resistance-training interventions (e.g. using weight machines that provide resistance to movement), and other interventions (e.g. Pilates). The average age of participants was 48.6 years. Trials were conducted in seven countries, and most studies (58.3%) included only female participants. Exercise trials ranged from 4 to 20 weeks. The stretching exercise programs ranged from 40 to 60 minutes, 1 to 3 times a day. The intensity of the stretches (e.g. how far the stretch was taken in the available range of motion) was not reported in most cases. The time each stretch was held ranged from 6 to 60 seconds. The targeted muscles were usually of both the upper and lower extremities, neck, and back. The flexibility training was either supervised or done at home. Our main comparison was flexibility exercise versus land-based aerobic training.

Key results at the end of treatment for our main comparison

Compared with land-based aerobic exercise training, flexibility exercise resulted in little benefit at 8 to 20 weeks' follow-up.

Each measure below was measured on a scale from 0 to 100, with lower scores better.

Health-related quality of life: People who received flexibility exercise training were 4% worse (ranging from 6% better to 14% worse).

• People who had flexibility training rated their quality of life as 46 points.

• People who had aerobic training rated their quality of life as 42 points.

Pain intensity: People who received flexibility exercise training were 5% worse (ranging from 1% better to 11% worse).

• People who had flexibility training rated their pain as 57 points.

• People who had aerobic training rated their pain as 52 points.

Fatigue: People who received flexibility exercise training were 4% better (ranging from 13% better to 5% worse).

• People who had flexibility training rated their fatigue as 67 points.

• People who had aerobic training rated their fatigue as 71 points.

Stiffness: People who received flexibility exercise training were 30% better (ranging from 8% better to 51% better).

• People who had flexibility training rated their stiffness as 49 points.

• People who had aerobic training rated their stiffness as 79 points.

Physical function: People who received flexibility exercise training were 6% worse (ranging from 4% better to 16% worse).

• People who had flexibility training rated their physical function as 23 points.

• People who had aerobic training rated their physical function as 17 points.

Withdrawal from treatment

A total of 18 per 100 people dropped out of the flexibility exercise training group for any reason compared to 19 per 100 people from the aerobic training group.

Harms

We found no clear information on harms. One study reported that one participant had swelling (tendinitis) of an ankle tendon (Achilles), but it is unclear if this was related to participation in the flexibility exercise.

Quality of evidence

The evidence does not show that flexibility exercise significantly improves health-related quality of life, pain, fatigue, or physical function. The number of people dropping out from each group was similar. Although the evidence suggests that flexibility exercise improves stiffness, caution is advised in interpretation of these results, as this improvement was seen in only one study with very few participants. We considered the overall certainty of the evidence to be low to very low due to study design issues, the small number of participants, and low certainty of results.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Flexibility exercise training compared with aerobic exercise training for adults with fibromyalgia

Flexibility exercise training compared with aerobic exercise training for adults with fibromyalgia

Patient or population: adults with fibromyalgia
Settings: group and home program
Intervention: flexibility exercise training
Comparison: aerobic training

Outcome: measured at the end of intervention

Outcomes	Anticipated absolute effects* (95% CI)		Relative effect (95% CI)	№ of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with aerobic (end of intervention)	Risk with flexibility				
Health-related quality of life assessed with: FIQ Total (0 is best) 0-to-100-millimeter scale Follow-up: range 12 weeks to 20 weeks[5]	Mean health-related quality of life was 42 mm.	Mean 4.14 mm higher (5.77 lower to 14.05 higher)	-	193 (2 RCTs)	⊕⊖⊖⊖ VERY LOW[1,2,3,4]	Absolute change was 4% worse (6% better to 14% worse). Relative change[7] in the flexibility groups compared to the aerobic groups was 7.53% worse (10.5% better to 25.5% worse). NNTB n/a[6]
Pain intensity assessed with: VAS (0 is best) 0-to-100-millimeter scale Follow-up: range 8 weeks to 20 weeks[8]	Mean pain intensity was 52 mm.	Mean 4.72 mm higher (1.39 lower to 10.83 higher)	-	266 (5 RCTs)	⊕⊖⊖⊖ VERY LOW[1,3,4]	Absolute change was 5% worse (1% better to 11% worse). Relative change in the flexibility groups compared to the aerobic groups was 6.7% worse (2% better to 15.4% worse).[7] NNTB n/a[6]
Fatigue assessed with: FIQ and SF-36 converted (0 is best) 0-to-100-millimeter scale Follow-up: range 8 weeks to 20 weeks[9]	Mean fatigue was 71 mm.	Mean 4.12 mm lower (13.31 lower to 5.06 higher)	-	75 (2 RCTs)	⊕⊖⊖⊖ VERY LOW[1,4]	Absolute change was 4% better (13% better to 5% worse). Relative change in the flexibility groups compared to the aerobic groups was 6.02% better (19.4% better to 7.4% worse).[7] NNTB n/a[6]
Stiffness assessed with: FIQ (0 is best) 0-to-100-millimeter scale Follow-up: 8 weeks[10]	Mean stiffness was 79 mm.	Mean 29.6 mm lower (51.47 lower to 7.73 lower)	-	15 (1 RCT)	⊕⊖⊖⊖ VERY LOW[4,11]	Absolute change was 30% better (8% better to 51% better). Relative change in the flexibility group compared to the aerobic

						group was 39% better (10% better to 68% better).[7] NNTB n/a[6]
Physical function assessed with: FIQ and SF-36 converted (0 is best) 0-to-100-millimeter scale Follow-up: range 8 weeks to 20 weeks[12]	Mean physical function 17 units.	Mean 6.04 units higher (3.95 lower to 16.03 higher)	-	60 (1 RCT)	⊕⊖⊖⊖ VERY LOW[1,4]	Absolute change was 6% worse (4% better to 16% worse). Relative change in the flexibility group compared to the aerobic group was 13.97% worse (9.1% better to 37.1% worse).[7] NNTB n/a[6]
Withdrawals All-cause attrition Follow-up: 8 to 20 weeks	Study population 19 per 100	18 per 100 (11 to 29)	RR 0.97 (0.61 to 1.55)	301 (5 RCTs)	-	Absolute change was 1% fewer withdrawals in the flexibility groups (8% fewer to 21% more). Relative change in the flexibility group was 3% fewer (39% fewer to 55% more).
Adverse events—increase in symptoms, injuries, or serious adverse events	Studies did not measure or report events.	Not all studies measured or reported events.	-	No reliable estimate	⊕⊖⊖⊖ VERY LOW[1,4]	In 1 of the 5 studies, 1 participant in the flexibility group was reported as having a minor adverse event. The following statement was provided: "a patient in the FLEX group had tendinitis of the Achilles tendon, which responded to treatment with local heat and a reduction in exercise for 14 days" (McCain 1988; page 1138). However, it is unclear whether the tendinitis was related to intervention participation.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CI: confidence interval; **FIQ:** Fibromyalgia Impact Questionnaire; **NNTB:** number needed to treat for an additional beneficial outcome; **NNTH:** number needed to treat for an additional harmful outcome; **RCT:** randomized controlled trial; **RR:** risk ratio; **SF-36:** 36-item Short Form Health Survey; **VAS:** visual analogue scale

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded two levels due to risk of bias (e.g. selection and performance bias).
[2]Downgraded one level due to inconsistency (i.e. heterogeneity among trials found).
[3]Downgraded two levels because flexibility was used as a proxy (i.e. flexibility exercise was used along with relaxation as the control in the study).
[4]Downgraded one level due to imprecision (sample size lower than 400 rule-of-thumb).

[5]Study authors: Richards 2002; Valim 2003.

[6]NNTB or NNTH was not calculated, as there were no clinically important between-group differences.

[7]Relative change calculation as per Cochrane Musculoskeletal Review Group procedures: absolute change divided by the baseline mean of the highest-weighted aerobic group. Richards 2002 (value was 55 on a 0-to-100-point scale on the FIQ for **health-related quality of life**, and 70.4 on a 0-to-100-point scale on the VAS for **pain**). Valim 2003 (value was 68.4 points on a 0-to-100-point scale on the SF-36 Vitality for **fatigue**, and 43.23 on a 0-to-100-point scale on the SF-36 for **function**). Bressan 2008 (value was 75.7 points on a 0-to-100-point scale on the FIQ for **stiffness**).

[8]Study authors: Bressan 2008; Matsutani 2012; McCain 1988; Richards 2002; Valim 2003.

[9]Study authors: Bressan 2008; Valim 2003.

[10]Study author: Bressan 2008.

[11]Downgraded one level for possible selection and performance bias.

[12]Study author: Valim 2003.

BACKGROUND

Description of the condition

Fibromyalgia syndrome is defined as a condition of generalized, chronic pain lasting at least three months accompanied with widespread muscular tenderness (Wolfe 2016). Individuals with this condition may also experience some degree of decreased energy, fatigue, stiffness, sleep disturbances, depression, memory problems, anxiety, tenderness to touch, balance challenges, and sensitivity to loud noises, bright lights, odors, and cold (Bennett 2014; Macfarlane 2017; Wolfe 2016). Additionally, cognitive impairment, sexual dysfunction, and reduced physical functioning may be experienced (Ghavidel-Parsa 2015; Zettel-Watson 2011). These symptoms compromise quality of life, thus impacting home and work environments and possibly leading to a loss of productivity, unemployment, and disability (Ghavidel-Parsa 2015). Genetic factors may contribute to the development of fibromyalgia through a dysfunctional stress response resulting from the hypothalamo-pituitary axis following a triggering event (Fitzcharles 2013).

Based on 2012 Canadian diagnostic criteria, available estimates of the prevalence of fibromyalgia in Canada suggest that 2% to 3% of the population experiences the condition, and that it more commonly affects females (Fitzcharles 2013). Other countries have reported similar prevalence rates, using Wolfe 1990 or Wolfe 2010 diagnostic criteria, ranging from 0.4% in Greece and 0.6% in Thailand to 6.4% in the United States and 8.8% in Turkey (Queiroz 2013). Worldwide, the estimated prevalence of fibromyalgia based on previous diagnostic criteria is 2.7%, including 4.1% females and 1.4% males (Queiroz 2013). Following the modified 2010 American College of Rheumatology diagnostic criteria for fibromyalgia, the prevalence of fibromyalgia in the United Kingdom has increased from 1.7% to 5.4% (Jones 2015). With these more recent criteria, the condition is still disproportionately experienced by females, though a greater proportion of males are now being diagnosed with fibromyalgia, as sex or gender ratios have reduced from 13.7:1 to 2.3:1 (Jones 2015). The most recent fibromyalgia criteria, updated in 2016, have identified 96.2% agreement with the 2011 criteria, suggesting that the increased diagnoses rates since the 2011 criteria may continue (Ablin 2017; Wolfe 2016). Fibromyalgia is present among individuals with musculoskeletal disorders, those with other illnesses such iHIV infection or Lyme disease (Buskila 1990; Dinerman 1992), and people with psychological disorders such as depression (MacFarlane 1999). This highlights the diversity of individuals who may experience this condition (Wolfe 2016), as well as the varying comorbidity present.

Many people with fibromyalgia are hesitant to engage in physical activity due to a fear of symptom exacerbation following exercise (Nijs 2013), thus potentially increasing risks of additional comorbidities (Nijs 2013). Individuals with fibromyalgia often experience comorbid illnesses, including musculoskeletal conditions, cardiovascular disorders, endocrinological disorders, spondylosis/intervertebral disc disorders and other back problems, irritable bowel syndrome, interstitial cystitis/painful bladder syndrome, chronic pelvic pain, temporomandibular joint disorder, depression, anxiety, and other psychiatric disorders (Ghavidel-Parsa 2015).

Fibromyalgia care and comorbidities require significant healthcare resources and costs (Ghavidel-Parsa 2015). Healthcare costs include healthcare visits and hospitalizations, pharmaceuticals, and extensive diagnostic testing (Ghavidel-Parsa 2015). On average, individuals with fibromyalgia make 10 to 18 primary care appointments per year and are hospitalized every 3 years (Ghavidel-Parsa 2015). Several pharmacotherapy treatments have shown tier 2 evidence for moderate pain relief (Macfarlane 2017). Cochrane Reviews of these therapies have included pregabalin and gabapentin (antiepileptics) (Derry 2016 Macfarlane 2017; Roskell 2011; Wiffen 2013), cyclobenzaprine (a muscle relaxant) (Macfarlane 2017; Tofferi 2004), duloxetine, milnacipran, and fluoxetine (serotonin and norepinephrine reuptake inhibitors) (Hauser 2012; Hauser 2013; Macfarlane 2017; Ormseth 2010; Roskell 2011), tramadol (an opioid pain medication and serotonin and norepinephrine reuptake inhibitor) (Macfarlane 2017; Roskell 2011), and amitriptyline (a tricyclic antidepressant) (Hauser 2012; Macfarlane 2017; Moore 2012) and the evidence has been of moderate and high certainty. Non-pharmacologic treatments of fibromyalgia have recently been recommended (Fitzcharles 2013; Macfarlane 2017). Cochrane Reviews of non-pharmacologic treatments have identified moderate-certainty evidence for fibromyalgia management including aerobic exercise (Bidonde 2017; Busch 2007). Additional reviews have identified low-certainty evidence for aquatic exercise (Bidonde 2014), resistance exercise (Busch 2013), cognitive behavioral therapy (Bernardy 2013), acupuncture (Deare 2013), and mind-body therapy (Theadom 2015).

Exercise training is now recognized as the cornerstone of treatment and management strategies for fibromyalgia as it represents the strongest evidence available (Fitzcharles 2013; Macfarlane 2017). Non-pharmacological treatments, especially exercise training, are recommended as the first treatment option for fibromyalgia (Macfarlane 2017). Fibromyalgia treatment recommendations include individualized exercise training tailored to a person's physical abilities and level of conditioning in exercises enjoyed or preferred by the individual (Fitzcharles 2013; Nijs 2013).

Description of the intervention

Flexibility exercise training is a type of exercise that focuses on improving or maintaining the range of motion in muscles and joint structures by holding or stretching the body in specific positions (ACSM 2013). Joint range of motion is an important physical characteristic that influences the capacity to perform activities of daily living (Mulholland 2001). Muscle stretching exercises increase the length of the muscle (or muscle group) beyond what would customarily be used in normal activity. This can improve non-clinical populations' range of motion temporarily right after flexibility exercises, as well as chronically after approximately three to four weeks of regular stretching at a frequency of at least two to three times a week (de Weijer 2003; Decoster 2005; Guissard 2006; Kokkonen 2007; Radford 2006; Reid 2004). Range of motion may improve in as few as 10 sessions with an intensive program (Guissard 2004).

Different types of stretching exercises can improve range of motion. Ballistic methods use the momentum of the moving body segment to produce the stretch. This is commonly used as warm-up (Woolstenhulme 2006). Dynamic or slow movement stretching involves a gradual transition from one body position to another, with a progressive increase in reach and range of motion as the movement is repeated several times (McMillian 2006). Static stretching involves slowly stretching a muscle-tendon

group and holding the position for a period (i.e. 10 s to 30 s for young people and 30 s to 60 s for older people) Decoster 2005; Feland 2001). Static stretching can be active or passive (Winters 2004). Active static stretching involves holding the stretched position using the strength of the agonist muscle. In passive static stretching, a position is assumed while holding a limb or other part of the body with or without the assistance of a partner or device. Static stretching, holding at the point of tightness or slight discomfort, is the most commonly used stretching mode (Kay 2015). Proprioceptive neuromuscular facilitation (PNF) methods take several forms but typically involve an isometric contraction of the selected muscle–tendon group followed by a static stretching of the same group and requires partner assistance (Rees 2007; Sharman 2006). Proprioceptive neuromuscular facilitation regularly produces greater increases in range of motion, however it can be problematic, as performing these contractions can be painful and induce muscle damage (Kay 2015).

Low levels of flexibility have been associated with postural problems, pain, injuries, decreased local vascularization, and increased neuromuscular tensions (Coelho 2008). In fact, flexibility training programs have been used to improve a person's well-being and as a tool for symptom management in different clinical populations such as those with major depressive disorders (Ambrose 2015; Costa 2009; Jones 2006; Lanuez 2011).

How the intervention might work

The main goal of flexibility training is usually to improve or maintain range of motion in major muscle–tendon groups in accordance with individualized goals (ACSM 2013; Garber 2011). Flexibility training improves postural stability and balance, Costa 2009, and enhances physical function, range of motion, Jones 2002; Valencia 2009, and muscle strength, Jones 2006. Flexibility training also decreases such fibromyalgia symptoms as pain, (Valencia 2009), muscle stiffness (Chen 2011), fatigue, and psychological factors (anxiety and depression) (Ambrose 2015; Lanuez 2011; Valencia 2009). It may be speculated that improved flexibility training could also enhance self-perceived ability to perform activities of daily living, and thereby improve psychosocial factors such as depressive symptoms, Soriano-Maldonado 2016, and social interaction, which are related to mental health and mood (Peluso 2005). Flexibility training may thus be beneficial for both fitness improvements and symptom control. Since stiffness and reduced range of motion have been shown to reduce health-related quality of life (HRQoL) in individuals with fibromyalgia (Valencia 2009), flexibility training may contribute to decreasing these physical difficulties thus improving HRQoL.

Flexibility training may be implemented as a program of static stretches that are held for 10 s to 30 s (ACSM 2013). Such activity may be used as part of relaxation programs that have demonstrated a positive effect on physical functioning and pain (Theadom 2015).

Why it is important to do this review

Flexibility exercises are advocated for the general public as a method to address stiffness and increase or maintain range of motion of major joints of the body (such as shoulders, hips, knees, ankles, back, neck) in order to maintain or improve general physical function (ACSM 2013). Since incorporating exercise into one's daily routine is not a small endeavour, it is the responsibility of clinicians and researchers to identify whether flexibility training should be undertaken both to improve and maintain physical function and to improve symptoms of fibromyalgia. If this form of exercise contributes to symptom improvement, it is important to identify which symptoms are most affected and the magnitude of the improvement. This review is important because flexibility training exercise is commonly recommended by consumer organizations designed to provide peer support (such as the National Fibromyalgia Association (www.fmaware.org/)). These organizations include individuals with fibromyalgia and healthcare providers, policymakers, and researchers (such as the National Fibromyalgia and Chronic Pain Association (https://fibroandpain.org/). This review was important to examine whether flexibility training does or does not have an effect on symptoms of fibromyalgia and HRQoL. Definitions for some of the terms utilized in this review can be found in the "Glossary of terms" (Appendix 1).

OBJECTIVES

- To evaluate the benefits and harms of flexibility exercise training interventions for adults with fibromyalgia.

To assess the following specific comparisons:

- Flexibility versus untreated controls (e.g. usual medical treatment)
- Flexibility versus aerobic interventions (e.g. treadmill walking)
- Flexibility versus resistance training (e.g. progressive training using weight machines)
- Flexibility versus other interventions (e.g. Pilates, friction massage, medication)

METHODS

Criteria for considering studies for this review

Types of studies

We included trials described as randomized, even if the methods of generating the random sequence were unclear or unreported, or the method of allocating participants was likely to be quasi-random (e.g. by alternation, date of birth, or similar pseudo-randomized method). We did not include studies using cross-over or cluster-randomized designs. We set no restriction on the number of participants included in the studies.

Types of participants

We included studies that examined adults with fibromyalgia (≥ 18 years of age). We selected studies that used published criteria for the diagnosis (or classification) of fibromyalgia. The American College of Rheumatology (ACR) 1990 criteria have long been used as the standard for classifying individuals as having fibromyalgia (Wolfe 1990). By this method, an individual is classified as having fibromyalgia when they have experienced widespread pain lasting longer than three months with at least 11 active tender points (TP). Tender points are noted at 18 designated locations on the body and are defined as active if pain can be elicited by applying 4-kilogram tactile pressure.

A diagnostic tool, ACR 2010 (Wolfe 2010), which does not rely upon a physical tender point examination, is also available both as a clinician-administered questionnaire and as a survey questionnaire (Wolfe 2011). This measure includes the Widespread Pain Index

(19 areas representing anterior and posterior axis and limbs), in addition to a Symptom Severity Scale that contains items related to secondary symptoms such as fatigue, sleep disturbances, cognition, and somatic complaints. Scores on both measures are used to determine whether a person qualifies for a "case definition" of fibromyalgia. This tool has been found to correctly classify 88% of cases that meet ACR 1990 criteria, and it allows ongoing monitoring of symptom change among individuals with a current or previous fibromyalgia diagnosis (Wolfe 2010). Although measures focusing on tender point counts have been widely applied in clinical and research settings, the methods described by Wolfe 2010 and Wolfe 2011 seem to classify people with fibromyalgia more efficiently, while allowing improved monitoring of disease status over time.

We also included studies where participants were diagnosed with fibromyalgia under different published diagnostic criteria, such as those by Smythe 1979 and Yunus 1981. Although some differences between published fibromyalgia diagnostic or classification criteria are known, for the purposes of this review, we considered all criteria to be acceptable and comparable.

Types of interventions

We examined trials that studied flexibility exercise training interventions regardless of frequency, duration, or intensity. We defined flexibility as movements of a joint or a series of joints through the complete range of motion that targeted major muscle-tendon units (ACSM 2013).

We have presented data on interventions using the Frequency, Intensity, Time, Type, Volume, Pattern and Progression (FITT-VP) principles of exercise prescription (Table 1) outlined for healthy individuals in Appendix 2 (ACSM 2013).

Comparator interventions included land-based aerobic training (e.g. treadmill walking), resistance training (e.g. progressive training using weight machines), and other interventions (e.g. Pilates, friction massage, Tai Chi, medication, aquatic biodanza). It should be noted that most aerobic and strength training interventions included brief (typically 5 to 10 minutes) warm-up and cool-down exercises before and after the main exercise component. These warm-up and cool-down components usually included a mix of stretching exercise and light aerobic exercise.

The main comparisons assessed in this review included the following.

- Flexibility exercise training versus untreated control
- Flexibility exercise training versus land-based aerobic exercise
- Flexbility exercise training versus resistance training
- Flexibility exercise training versus other interventions

For the purposes of this review, we were interested in interventions in which the effects of flexibility exercise training could be isolated, therefore we excluded studies that combined flexibility exercise training with other interventions or education.

Types of outcome measures

Major outcomes

Seven outcomes were designated as major outcomes: HRQoL, pain intensity, fatigue, stiffness, physical function, adverse events, and number of participants who withdrew or dropped out. Three outcomes were designated as minor outcomes: tenderness, depression, and greater than 30% improvement in pain. In selecting these outcomes, we sought the opinion of consumers involved in the team and considered the consensus statement of Choy 2009a regarding a core set of outcome measures for clinical trials in fibromyalgia as anticipated effects of flexibility exercise training on physical fitness. We extracted data for selected outcomes at baseline, end of intervention (post-treatment), and follow-up data. Review criteria required each included study to report measurement of one or more outcomes for at least one of these time periods.

When an included study used more than one instrument to measure a particular outcome, we applied the following preferred hierarchy to choose the outcome measure for analysis.

- Health-related of life. This outcome consists of multidimensional indices used to measure general health status or HRQoL, or both (Choy 2009a). When included studies used more than one instrument to measure HRQoL, we preferentially extracted data from the Fibromyalgia Impact Questionnaire (FIQ Total; Bennett 2009; Burckhardt 1991), followed by the Short Form Health Survey questionnaire (either the SF-36 total or the SF-12 total; Busija 2011; Ware 1993) and the EuroQol-5D (EQ-5D) (Wolfe 1997).
- Pain intensity. The International Association for the Study of Pain defined pain as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (Merskey 1994). For the purposes of this review, we focused on one aspect of the pain experience, i.e. pain intensity. When more than one measure of pain intensity was reported in a single study, we preferentially extracted measures of average pain intensity (as opposed to worst, least, or current pain) assessed by visual analogue scale (VAS), FIQ Pain, McGill pain VAS followed by the Numerical Pain Rating Scale. In studies where unidimensional measures of pain intensity were not reported, we extracted composite measures that included pain intensity and interference (SF-36 or Rand 36 Bodily Pain Scale; Ware 1993) or pain intensity and suffering from pain (Multidimensional Pain Inventory - Pain Severity scale).
- Fatigue. Fatigue is recognized by individuals with fibromyalgia and clinicians alike as an important symptom (Choy 2009a). Fatigue can be measured in a global manner, in which an individual rates fatigue on a single-item scale, or using a multidimensional tool that breaks the experience of fatigue down into two or more dimensions such as general fatigue, physical fatigue, mental fatigue, reduced motivation, reduced activity, and degree of interference with activities of daily living (Boomershine 2012). We accepted both uni- and multidimensional measures for this outcome. When included studies used more than one instrument to measure fatigue, we preferentially extracted the fatigue VAS (FIQ Fatigue, or single-item fatigue VAS), followed by the SF-36 or Rand 36 Vitality subscale, the Chalder Fatigue Scale (total), the Fatigue Severity Scale, and the Multidimensional Fatigue Inventory.
- Stiffness. In focus groups conducted by Arnold 2008, individuals with fibromyalgia "... remarked that their muscles were constantly tense. Participants alternately described feeling as if their muscles were 'lead jelly' or 'lead Jell-O,' and this resulted in a general inability to move with ease and a feeling of stiffness."

We used a common measure of stiffness encountered in this literature, i.e. the FIQ stiffness subscale.

- Physical function. This outcome focuses on the basic actions and complex activities considered "essential for maintaining independence, and those considered discretionary that are not required for independent living, but may have an impact on quality of life" (Painter 1999). Since cardiorespiratory fitness, neuromuscular attributes (e.g. muscular strength, endurance, and power), and muscle and joint flexibility are important determinants of physical function, this outcome is highly relevant as an outcome of exercise interventions. When more than one measure of physical function was available within a study, we preferentially extracted data for the FIQ physical impairment scale (Burckhardt 1991), followed by the Health Assessment Questionnaire disability scale (HAQ), the SF-36 or Rand 36 Physical Function, the Sickness Impact Profile – Physical Disability (Bergner 1981), and the Multidimensional Pain Inventory household chores scale (Huskisson 1976; Huskisson 1983).
- Adverse events. We extracted the number of participants who experienced adverse events during the intervention (i.e. injuries, exacerbations of pain and/or other fibromyalgia symptoms). If this information was not available, we extracted the nature of the adverse events in a narrative report.
- Withdrawals. We reported the number of participants who withdrew or dropped out of the study for any reason.

Minor outcomes

The following is a rationale and preference listing of minor outcomes. Among the three outcomes designated as minor outcomes, we have included one psychological and one physical variable that could potentially improve with flexibility exercise training.

- Depression. This is a common mental disorder characterized by depressed mood, loss of interest or pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, low energy, and poor concentration. These problems can become chronic or recurrent and lead to substantial impairments in a person's ability to attend to his or her everyday responsibilities (WHO 2017). In focus groups conducted by Arnold 2008, the emotional disturbances most commonly experienced by participants with fibromyalgia included depression and anxiety. A complete understanding of depression and how best to assess it in fibromyalgia trials is still uncertain and is an active research issue (Mease 2009). However, because people with significant depression are commonly excluded from fibromyalgia intervention studies, the discriminatory power of these instruments is underestimated (Choy 2009b). We preferentially extracted Beck Depression Inventory (BDI) Cognitive/Affective subscale scores followed by BDI total, BDI without fibromyalgia symptoms; Short Form translated SF-36; Hamilton Depression Scale; Center for Epidemiologic Studies Depression Scale (CES-D); Fibromyalgia Impact Questionnaire (FIQ) FIQ translated- depression subscale; Mental Health Inventory (MHI) depression subscale; Arthritis Impact Measurement scales (AIMS) - depression subscale; Hospital Anxiety and Depression Scale - depression (HADS); Symptom Checklist 90 (SCL-90-R) - depression; and the Psychological General Well-Being (PGWB depression score).
- Tenderness. Tenderness was defined as discomfort produced as an evoked response to mechanical pressure (Dadabhoy 2008; Gracely 2003). Although there are concerns that measures of tenderness can be biased by cognitive and emotional aspects of pain perception, many studies have supported the utility of measurement of tenderness in fibromyalgia using either TP counts or pain pressure threshold (Dadabhoy 2008). A TP is identified when pressure of 4 kg is perceived as painful. When included studies used more than one instrument to measure tenderness, we preferentially extracted the TP count followed by pain pressure threshold (dolorimetry score, based on at least six of the 18 ACR TPs) and the total myalgic score (sum/mean of ordinal rating of response to thumb pressure across 18 TPs).
- Improvement greater than 30% in pain. A 30% reduction is considered a benchmark for a moderately important change in pain intensity, and the consensus group Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommends this measure for interpreting clinical trial efficacy (Dworkin 2008). We extracted data on the number of participants who met this criterion for intervention efficacy when this information was available.

Search methods for identification of studies

The team Information Specialist conducted a comprehensive search in nine databases for studies of physical activity interventions in adults with fibromyalgia. The citations found in the electronic and manual searches were screened and then classified by the type of exercise training. This comprehensive search captured all types of physical activity intervention studies, of which only the subset classified as studies of flexibility training interventions was included in this review.

Electronic searches

We searched the following databases from database inception to 31st of December, 2017 using the methods outlined in Chapter 6 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). We applied no language restrictions. The full search strategies for each database are shown in the Appendices as indicated below.

- MEDLINE (Ovid) MEDLINE In-Process and MEDLINE 1946 to 31st of December 2017 (Appendix 3)
- Embase (Ovid) Embase Classic + Embase 1947 to 31st of December 2017 (Appendix 4)
- CINAHL (EBSCO) (Cumulative Index to Nursing and Allied Health Literature) 1982 to 31st of December 2017 (Appendix 5)
- Cochrane Library (Wiley) 2003, Issue 1 to present (Appendix 6)
  - Cochrane Database of Systematic Reviews (Cochrane Reviews)
  - Database of Abstracts of Reviews of Effects (DARE)
  - Cochrane Central Register of Controlled Trials (CENTRAL)
  - Health Technology Assessment Database (HTA)
  - NHS Economic Evaluation Database (EED)
- AMED (Ovid) (Allied and Complementary Medicine Database) 1985 to 31st of December 2017 (Appendix 7)
- Thesis and Dissertation Abstracts (ProQuest) 1743 to December 2017 (Appendix 8)
- PEDro (Physiotherapy Evidence Database) 1929 to December 2017 (Appendix 9)

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (www.clinicaltrials.gov/) 2000 to 31st of December 2017 (Appendix 10)
- World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (www.who.int/ictrp/en/) 2007 to 31st of December 2017 (Appendix 11)

Searching other resources

Two review authors independently reviewed reference lists from key journals; identified articles and reviews of all types of treatment for fibromyalgia; scrutinized all promising or potential references; and added appropriate titles to the search results.

Data collection and analysis

Review authors

The review authors are members of the Cochrane Musculoskeletal Group (CMSG) - Physical Activity and Fibromyalgia Team (see Acknowledgements). The review authors were trained in data extraction and 'Risk of bias' assessment using a standardized orientation program. They worked independently and in pairs with at least one physical therapist in each pair to extract data. Two additional members, our team consumers, assisted at several stages of the review. They were involved in selecting the outcomes, writing the Plain language summary, and reading the final draft for content and readability. The entire team met regularly to discuss progress, clarify procedures, make decisions regarding inclusion or exclusion of studies and classification of outcome variables, and work collaboratively in the production of this review.

Selection of studies

Two review authors independently examined the titles and abstracts of studies generated from the searches using a set of criteria (Appendix 12). The team used Covidence software to assist with independent screening of literature (Covidence 2015). We retrieved the full-text publications for all potentially relevant abstracts. All non-English reports were translated (Amanollahi 2013; López-Rodríguez 2012; Matsutani 2012). We then examined the full-text reports to determine study eligibility based on the selection criteria. Disagreements between the two review authors and questions regarding interpretation of inclusion criteria were resolved by discussion or by consulting a third review author if needed.

In keeping with Rosenthal's recommendations (Rosenthal 1995), publications referring to the same primary study (what we called 'companions') but presenting follow-up data in consequent publications were linked and presented as one. Likewise, published studies for which protocols were found in trial registries or were published were considered companions and presented as one.

Data extraction and management

We used electronic data extraction forms developed, piloted, and refined in our previous reviews to facilitate independent data extraction and consensus (Busch 2008). Pairs of review authors independently extracted the data. Any disagreements were resolved by consensus or involving a third person (AJB) if necessary. Two review authors (SYK, AJB) transferred data into the Review Manager 5 software file (RevMan 2014). We double-checked that data were entered correctly by comparing the data presented in the software with the study reports. We noted in the Characteristics of included studies table whether outcome data were not reported in a usable way (Assumpção 2017); instances when the data were obtained directly from study authors (Altan 2009; Assumpção 2017; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Richards 2002); and when data were transformed or estimated from a graph (Calandre 2009). If both unadjusted and adjusted values for the same outcome were reported, we extracted the adjusted values. If the data were analyzed based on an intention-to-treat (ITT) sample and another sample (e.g. per-protocol, as-treated), we extracted the ITT data. Due to changes in the methods (e.g. risk of bias), we reassessed studies included in the previous review, (Busch 2002; Busch 2007), for this updated review.

We extracted the following data from the included studies.

- Methods: study design, total duration of study and follow-up (if applicable), and date of study.
- Participants: N, n, mean age, age range, gender ratio, disease duration, diagnostic criteria, inclusion and exclusion criteria.
- Interventions, comparison, concomitant treatments recording:
  * for all interventions with an exercise component: frequency, duration of exercise sessions, intensity, mode, and congruence with American College of Sports Medicine (ACSM) guidelines for healthy adults (ACSM 2013);
  * for interventions with a non-exercise component: frequency, duration, and main characteristics.
- Outcomes: major and minor outcomes as indicated previously; additional outcomes assessed (recorded in the Characteristics of included studies table); means and standard deviations for tests at baseline and end of intervention (post-treatment) and follow-up for continuous outcomes.
- Characteristics of trial design as outlined in the Assessment of risk of bias in included studies section.
- Country of study, language of article, records of author contacts, trials registry record or protocol, and notable declarations of interest (recorded in the Characteristics of included studies table).

Assessment of risk of bias in included studies

We assessed risk of bias of studies based on the procedures recommended in the *Cochrane Handbook for Systematic Reviews of Interventions*. Two review authors independently evaluated the risk of bias in each included study using a customized form based on the Cochrane 'Risk of bias' tool (Higgins 2011). The tool addresses seven specific domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete outcome data, selective outcome reporting (including publication bias), and other sources of bias. For other sources of bias, we considered issues such as baseline inequities despite randomization.

We assessed each criterion as low, high, or unclear risk of bias according to the information provided in the studies and at times based on study author responses (Altan 2009; Assumpção 2017; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Richards 2002). We classified studies as having a low risk of bias if all key domains had low risk of bias and no serious flaws. We judged studies for which the absence of information or ambiguities prevented a determination of the potential for bias as at unclear risk of bias. In such cases, we revised our assessment if the authors responded to our requests for more information. Any disagreements between

the review authors were resolved through discussion at consensus meetings. If agreement could not be reached, involvement of a third review member was sought.

Measures of treatment effect

For continuous data, we used the group post-treatment means and standard deviations to calculate the effect sizes, employing Review Manager 5 software (RevMan 2014). We expressed effect sizes preferentially in the form of mean differences (MD) and 95% confidence intervals (95% CI). For dichotomous data, we used risk ratios (RR) and 95% CI.

We used Review Manager 5 software to generate forest plots to display the results (RevMan 2014). We used data from the latest follow-up assessments when evaluating long-term effects.

In the comments column of the Summary of findings for the main comparison, we provided the relative change and the number needed to treat for an additional beneficial outcome (NNTB). The NNTB was provided only when the outcome showed a clinically important difference. We calculated the NNTB for continuous measures using the Wells calculator (available at the CMSG Editorial office). For dichotomous outcomes, such as adverse events, we planned to calculate the NNTB from the untreated control group event rate and the risk ratio using the Visual Rx NNTB calculator. Data were not available, and we were unable to calculate the NNTB for dichotomous outcomes.

In accordance with the Philadelphia Panel 2001, we assumed a minimal clinically important difference (MCID) of 15 points on a 100-point continuous pain scale and a relative difference of 15% on all functional scales as being clinically relevant. The MCID was used in the calculation of NNTB for continuous outcomes. For dichotomous outcomes, the absolute risk difference was calculated using the risk difference statistic in Review Manager 5, with the result expressed as a percentage (RevMan 2014). For continuous outcomes, the absolute benefit was calculated as the improvement in the intervention group minus the improvement in the untreated control group, in the original units. Relative change calculation as per CMSG procedures: absolute change divided by the baseline mean (of the most weighted study) of the comparator groups.

Unit of analysis issues

We included studies with two or more parallel groups and examined any relevant comparison that allowed the evaluation of the effects of flexibility exercise training interventions on individuals with fibromyalgia. For example, a three-arm trial comparing flexibility versus drug treatment versus friction massage could appear in two separate analyses: flexibility versus medications, and flexibility versus friction massage. For details see the Characteristics of included studies table.

Dealing with missing data

When numerical data were missing, we contacted the author requesting the additional data required for analysis. We used open-ended questions to obtain the information needed to assess risk of bias and for the treatment effect. When numerical data were available only in graphic form, we used Engauge Digitizer version 5.1 to extrapolate means and standard deviations by digitalizing data points on the graphs (Mitchell 2012).

For dichotomous outcomes (e.g. number of withdrawals), we calculated the withdrawal rate using the number of participants randomized in the group as the denominator. For continuous outcomes (e.g. post-treatment in pain score), we calculated the MD or standardized mean difference (SMD) based on the number of individuals analyzed at that time point. When the number of individuals analyzed was not presented for each time point, we used the number of individuals randomized to each group at baseline. When means were not reported, we used medians.

When post-treatment standard deviations were unavailable, we used the standard deviations of the pre-test scores as estimates. When the variance was expressed using statistics other than standard deviation (e.g. standard error, confidence interval, P value), we computed standard deviations according to the methods recommended in Chapter 7 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). When we were unable to derive missing standard deviations using the above methods, we would impute them from other studies in the meta-analysis; however, this was not necessary for this review.

Assessment of heterogeneity

We assessed clinical and methodological diversity in terms of participants, interventions, outcomes, and study characteristics for the included studies to determine whether a meta-analysis was appropriate. We did this by reviewing data obtained from data extraction tables. We assessed heterogeneity through visual inspection of the forest plot to assess for obvious differences in result between the studies, and through the use of I^2 and Chi^2 statistical tests. As recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2017), we interpreted I^2 values as follows:

- 0% to 40%: might not be important;
- 30% to 60%: moderate heterogeneity;
- 50% to 90%: substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

We interpreted the Chi^2 test with a P value ≤ 0.10 as indicating statistical heterogeneity.

When we removed a trial from the analysis, we noted changes in both heterogeneity and effect size. Because I^2 involves overlapping categories (e.g. 0% to 40%, 30% to 60%), or 'ambiguous' zones, we explored statistical heterogeneity thoroughly when noted (e.g. I^2 between 50% and 60%). Given that values between 50% and 60% fall into an ambiguous zone, if we could find no apparent causes of heterogeneity, we kept the trial in the analysis and documented our decision.

Assessment of reporting biases

We planned to draw contour-enhanced funnel plots for each meta-analysis to assess publication reporting bias if a large enough sample of studies (i.e. more than 10 studies) was available and included in the meta-analysis (Sterne 2017).

If the randomized controlled trial (RCT) protocol was available, we compared the outcomes in the RCT protocol versus the outcomes in the published report. For studies published after 1 July 2005, we searched the WHO ICTRP and ClinicalTrials.gov for the RCT protocol.

We compared the fixed-effect estimate against the random-effects model to assess the possible presence of small-sample bias (i.e. by which intervention effect is more beneficial in smaller studies) in the published literature. In the presence of small-sample bias, the random-effects estimate of the intervention is more beneficial than the fixed-effect estimate (Sterne 2017).

Data synthesis

When two or more studies reported the same outcome and interventions were deemed sufficiently homogeneous, we pooled the data (meta-analysis) using Review Manager 5 (RevMan 2014). Before pooling data, we ensured the directionality of the data that permitted pooling; we arithmetically reversed selected scales as needed so higher values consistently had the same meaning. We ensured that scaling factors were consistent to permit calculations of MD (e.g. 10-centimeter scales expressed in millimeters to match 100-millimeter scales). We presented results grouped by common comparator, for example flexibility versus aerobics, flexibility versus resistance training, and flexibility versus other comparators. We included all studies for adverse events and withdrawals.

'Summary of findings' table

We used the GRADEpro software (GRADEpro GDT 2015) to prepare the 'Summary of findings' table for major outcomes for flexibility exercise training versus land-based aerobic training. In the 'Summary of findings' table, we integrated analysis of the certainty of the evidence and magnitude of effect of the interventions. We used the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias) to assess the certainty of the body of evidence at one of four levels, as follows.

- High certainty: further research is very unlikely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- Moderate certainty: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- Low certainty: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- Very low certainty: research shows substantial uncertainty about the estimate.

We downgraded the overall rating of the certainty of the evidence for the study (outcome by outcome) by at least one grade (using GRADE considerations) if the study had high or unclear risk of bias in a least one domain. We assigned GRADE certainty ratings separately for the seven major outcomes. Because of the comprehensive nature of the outcome variable of HRQoL, we gave it primacy over all other variables in the 'Summary of findings' table and the Plain language summary.

Subgroup analysis and investigation of heterogeneity

There were insufficient studies to conduct subgroup analysis as indicated in the review protocol (Busch 2015).

We assessed statistical heterogeneity among the trials using the heterogeneity statistics (Chi^2 test and I^2 statistic). We considered P values < 0.01 or I^2 > 50% to be indicative of significant heterogeneity. In the case of P value < 0.01 or I^2 > 50% (or both), we used a random-effects model instead of the fixed-effect model for meta-analysis. In addition, in the case of statistical heterogeneity, we scrutinized the studies for sources of clinical heterogeneity and methodological differences.

Sensitivity analysis

We planned to perform sensitivity analyses to investigate the impact of statistical heterogeneity and methodological weakness (i.e. high or unclear risk of selection bias and detection bias, or attrition rates > 20%).

RESULTS

Description of studies

See Characteristics of included studies; Characteristics of excluded studies; and Characteristics of studies awaiting classification

Results of the search

Our searches identified total of 6530 records. After removal of 2771 duplicates, 3759 records remained. We excluded 3478 records based on citation and abstract screening. We assessed 255 full-text articles, 1 thesis, and 25 trial registry records for eligibility. We excluded 96 full-text articles and 1 trial registry record. After assessing full-text physical activity articles for the type of intervention, we excluded 140 articles, 5 published study protocols, and 22 trial registry records because the intervention type was not flexibility. We included 14 full-text publications (12 primary studies and 2 companion papers), 1 thesis, and 2 trial registry records. For details see Figure 1.

Figure 1. Study flow diagram.

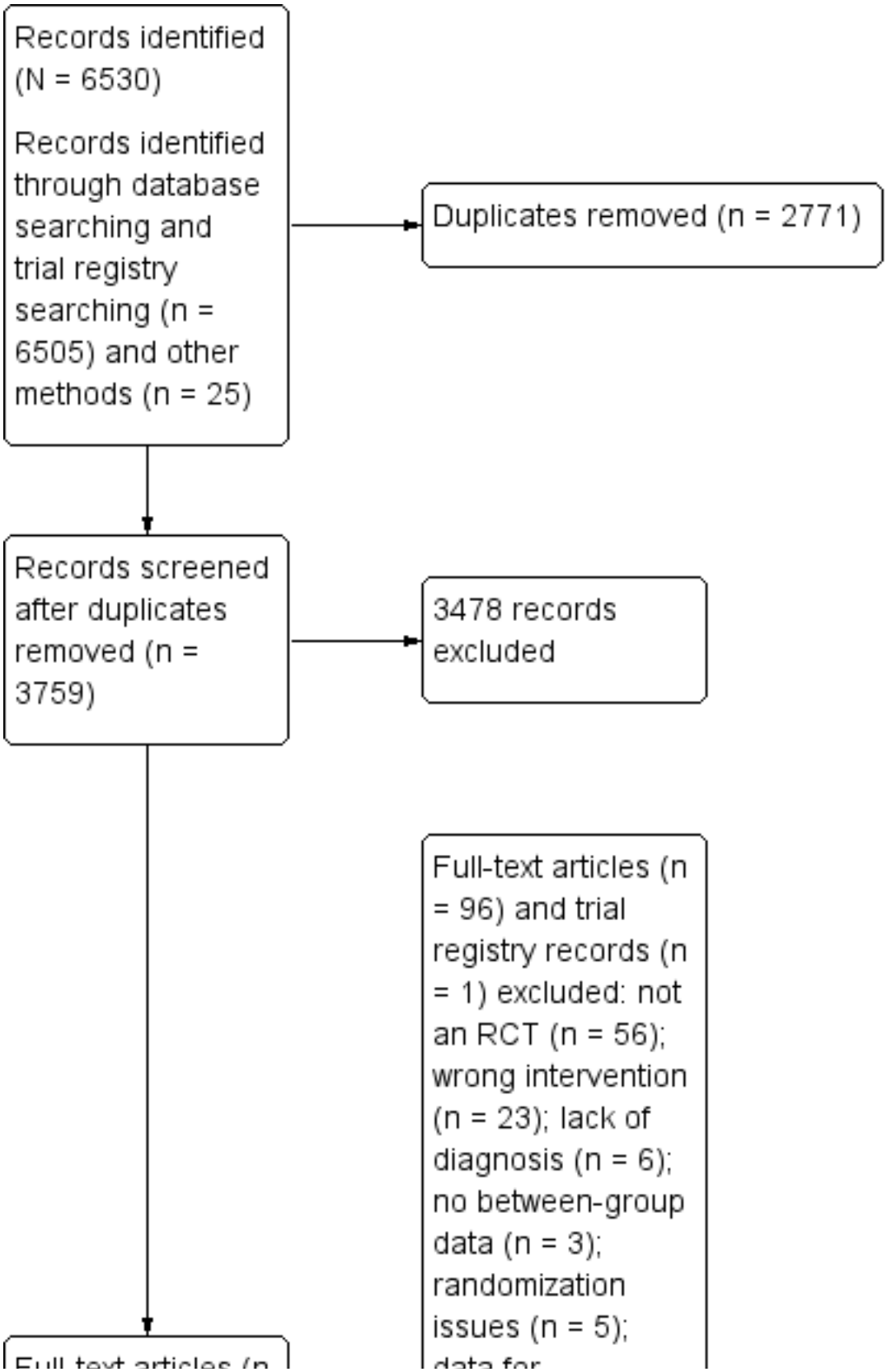
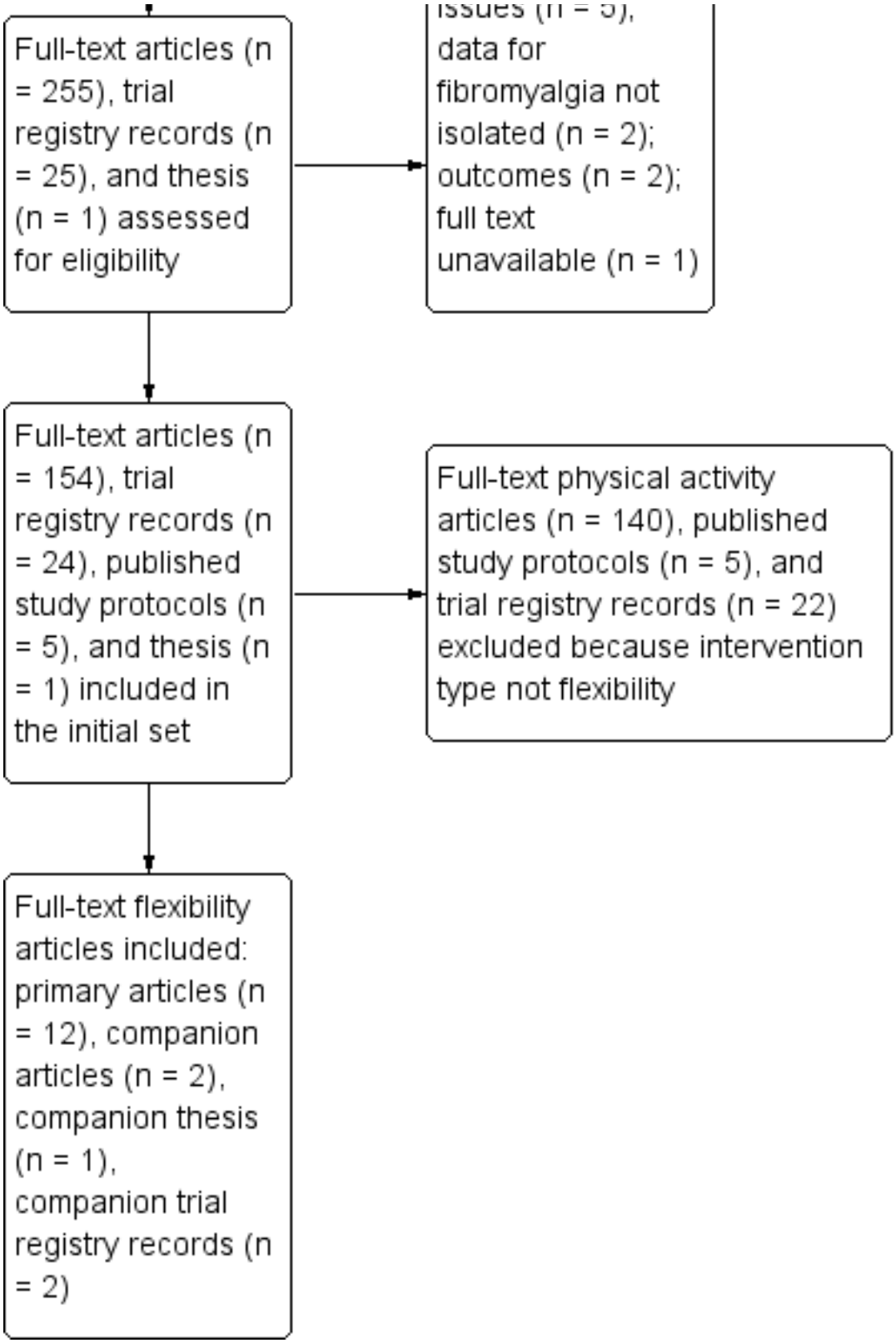


Figure 1. (Continued)



Included studies

Fourteen full-text reports, 2 registry records, and 1 thesis describing 12 unique flexibility exercise training studies met our selection criteria and were considered in this review (Altan 2009; Amanollahi 2013; Assumpção 2017; Bressan 2008; Calandre 2009; Gavi 2014; Jones 2002; López-Rodríguez 2012; Matsutani 2012; McCain 1988;

Richards 2002; Valim 2003). We used the two registry records (hereafter described as 'RCT protocols') to assess the certainty of studies.

The included studies were published between 1988 and 2017 and were conducted in seven different countries, as follows: Canada (1 study), the United States (1 study), Turkey (1 study), Brazil (4 studies), Iran (1 study), Spain (3 studies), and the United Kingdom (1 study). Nine of the 12 studies were published in English, with the remaining published in Spanish (López-Rodríguez 2012), Farsi (Amanollahi 2013), and Portuguese (Matsutani 2012). We contacted the authors of seven studies to request additional information needed to assess risk of bias, exercise intervention, and/or treatment effect (Altan 2009; Amanollahi 2013; Assumpção 2017; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Richards 2002). We received responses from the authors of six studies (Altan 2009; Assumpção 2017; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Richards 2002), and no response from the author of Amanollahi 2013. The outcomes extracted for all included studies are presented in Table 2.

Two studies had more than two study arms (Amanollahi 2013; Assumpção 2017). For details see the Characteristics of included studies section.

Participants

This review included 743 participants. Seven studies included only female participants (n = 448); one study included both male and female participants (Calandre 2009); and four studies did not specify the gender of participants (López-Rodríguez 2012; Matsutani 2012; McCain 1988; Richards 2002). The average duration of disease or symptoms since diagnosis ranged from 3 to 10 years. Nine studies did not report this information (Altan 2009; Amanollahi 2013; Assumpção 2017; Bressan 2008; Gavi 2014; López-Rodríguez 2012; Matsutani 2012; McCain 1988; Valim 2003). Based on 11 studies that provided mean ages and ranges, the average age of participants was 48.6, ranging from 35.8 to 56 (Richards 2002 did not provide mean ages, only median).

Fibromyalgia diagnosis was based on ACR 1990 criteria, Wolfe 1990, in all but one study (McCain 1988), where participants had to fulfill the diagnostic criteria of Smythe 1979.

The inclusion criteria for the studies included: age; diagnosis of fibromyalgia; willingness to keep their pharmacological treatment constant during the study period and not start new exercise or alternative therapies; being a patient of the study's health center; able to understand the procedures and follow the basic orientation given; pass the treadmill stress test (used to determine the effects of exercise on the heart; electrical activity of the heart is monitored during the test); provide consent; being sedentary women; never previously treated for fibromyalgia; newly diagnosed with fibromyalgia.

The exclusion criteria for the studies included: presence of an accompanying rheumatoid disease; unstable hypertension; severe cardiopulmonary problems; psychiatric disorders affecting participant compliance; infection; fever; severe physical impairment; inflammatory disease, uncontrolled endocrine diseases; allergic diseases (including allergy to chlorine); pregnancy; malignancy; inadequate cognitive level to understand the orientations and procedures; those who had never attended a swimming pool; had any disease susceptible to worsening with warm-water exercise; respiratory, metabolic, and rheumatic disease that could limit exercise; disease associated with autonomic dysfunction (e.g. arterial hypertension, diabetes); use of medications such as moderate or high dose of beta blockers, calcium channel blockers, antihypertensive, anticonvulsant, non-tricyclic antidepressants, and opioid analgesics; exercise within the last three months or current participation in a regular exercise program; inability to understand questionnaires; positive treadmill test (e.g. abnormal heart activity detected during exercise on the treadmill or myocardial ischemia detected); receipt of social security benefits; neurological or renal disease that would preclude involvement in an exercise program; current cigarette smoking; score ≥ 29 on the Beck Depression Inventory modified for fibromyalgia; missing 14 or more sessions or change in pharmacological treatment during the study; history or suspicion of neoplasia; amitriptyline within previous three months; ischemic heart disease; symptomatic cardiac arrhythmias; exercise-induced asthma; individuals for whom an alternative medical diagnosis could explain current symptoms; inability to attend classes; inability to co-operate; body mass index > 35; hyperthyroidism.

Interventions

Descriptions of trial interventions, including congruence with the ACSM criteria for flexibility in healthy adults (ACSM 2013), are detailed in the Characteristics of included studies section and in Table 1, Table 3, and Table 4 .

- Flexibility versus untreated controls (1 study). There was only one study in this comparison (Assumpção 2017). Exercise frequency was two times a week. The duration of the intervention was 12 weeks. The intensity was described as "stretch was gradually increased to point of moderate discomfort." The duration of each stretch was 30 seconds. Static stretches were used and targeted large muscles of upper and lower body. The flexibility intervention was 40 minutes in total. Volume (estimated from duration of stretch and repetitions that gradually increased from three to five through the intervention) ranged from 90 seconds to 2.5 minutes. The program was supervised.
- Flexibility versus aerobic training (5 studies). Exercise frequency ranged from one to three times a week: one time per week in Bressan 2008 and Matsutani 2012; two times per week in Richards 2002; and three times per week in McCain 1988 and Valim 2003. Duration varied from eight weeks, in Bressan 2008 and Matsutani 2012, to 20 weeks, in McCain 1988 and Valim 2003. None of the studies specified the intensity of the stretching exercises, therefore we were unable to determine if the stretches were taken to the intensity recommended by ACSM 2013, i.e. the point of feeling tightness or slight discomfort (ACSM 2013). The flexibility intervention time ranged from 40 to 60 minutes, with the average duration for each stretch 30 seconds. Two studies did not provide information on the duration of each stretch (McCain 1988; Richards 2002). Studies used static stretches for the major muscle-tendon units of both the upper and lower limbs, however in one study, McCain 1988, it was difficult to judge the type of stretches and body region (e.g. "Exercise consisted of flexibility maneuvers such that sustained heart rate responses greater than 115 beats per minute were not attained"). None of the studies outlined volume of stretches (i.e. total stretching time for each flexibility exercise), but we

could calculate the volume from the duration of stretch and repetitions in two studies (Bressan 2008; Matsutani 2012). Only two studies provided information on the number of repetitions for each stretching exercise, which ranged from four, in Bressan 2008, to five, in Matsutani 2012. Sessions were supervised in two studies (McCain 1988; Richards 2002). It was unclear if the stretching intervention was supervised in Bressan 2008, Matsutani 2012, and McCain 1988.

- Flexibility versus resistance exercise training (3 studies). Exercise frequency was two times a week for all three studies (Assumpção 2017; Gavi 2014; Jones 2002). The duration of the intervention ranged from 12 to 20 weeks. Intensity of the intervention was specified in only one study (Assumpção 2017), which stated that "the stretch intensity was increased gradually to the point of moderate discomfort." The flexibility intervention ranged from 40 to 60 minutes. The duration of each stretch was 30 seconds in Assumpção 2017 and Gavi 2014. The duration of each stretch was 60 seconds in Jones 2002. All studies used static stretches of major muscle-tendon units of both the upper and lower limbs. We could estimate the volume of stretches in one study (Assumpção 2017), as detailed above in the first bullet. The other studies did not provide volume of stretches and number of repetitions for each stretch. Flexibility exercise interventions were supervised in two studies (Assumpção 2017; Jones 2002). It was unclear if sessions were supervised in Gavi 2014.
- Flexibility versus other comparators (4 studies, 1 with 3 parallel arms). The frequency was three times per week in all but one study (López-Rodríguez 2012), where the frequency was two times per week. The duration of the intervention ranged from 4 to 12 weeks. None of the studies specified the intensity of the intervention. The mean length of the flexibility intervention was 60 minutes. One study did not specify the length of the flexibility intervention (Amanollahi 2013). The length of each stretch ranged from 6 to 30 seconds. Two studies did not specify the length of each stretch (Calandre 2009; López-Rodríguez 2012). For one study, we could calculate the volume of each stretch from the number of repetitions and length of each stretch: for example, the volume of each stretch was 90 seconds based on the 3 repetitions and 30 seconds hold per stretch in Amanollahi 2013. We could not calculate the volume in the remaining studies, as the number of repetitions or the length of each stretch (or both) was not provided.

Excluded studies

We excluded 3478 records that did not meet our inclusion criteria based on title and abstract screening (Figure 1). We examined 255 full-text articles, 1 thesis, and 25 trial registry records for possible inclusion in the review. We excluded full-text articles (n = 96) and trial registry records (n = 1) due to unmet criteria as follows: study design (n = 56); intervention (n = 23); diagnosis (n = 6); between-group data (n = 3); implementation of randomization (n = 5); isolation of data for fibromyalgia (n = 2); lack of designated outcomes (n = 2). The remaining 159 full-text articles (5 of which were published study protocols), 1 thesis, and 24 trial registry records represented RCTs examining effects of physical activity interventions for fibromyalgia. Of these, we ruled out 167 because the physical intervention did not have any flexibility-only intervention or the study was reviewed, or was designated to be reviewed, in another Cochrane Review in this series.

Risk of bias in included studies

'Risk of bias' assessments for the 12 included studies are provided in the 'Risk of bias' table in the Characteristics of included studies section and in Figure 2 and Figure 3. 'Risk of bias' assessments were based on primary article, protocol when available, and data supplemented by study author responses.

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

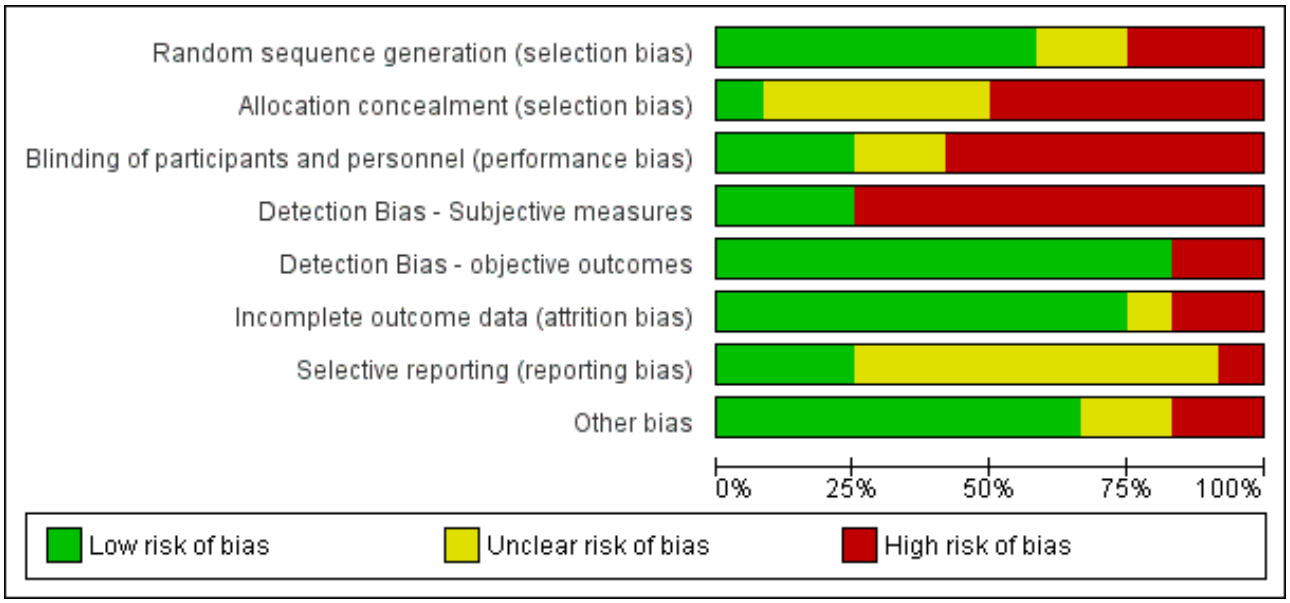


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

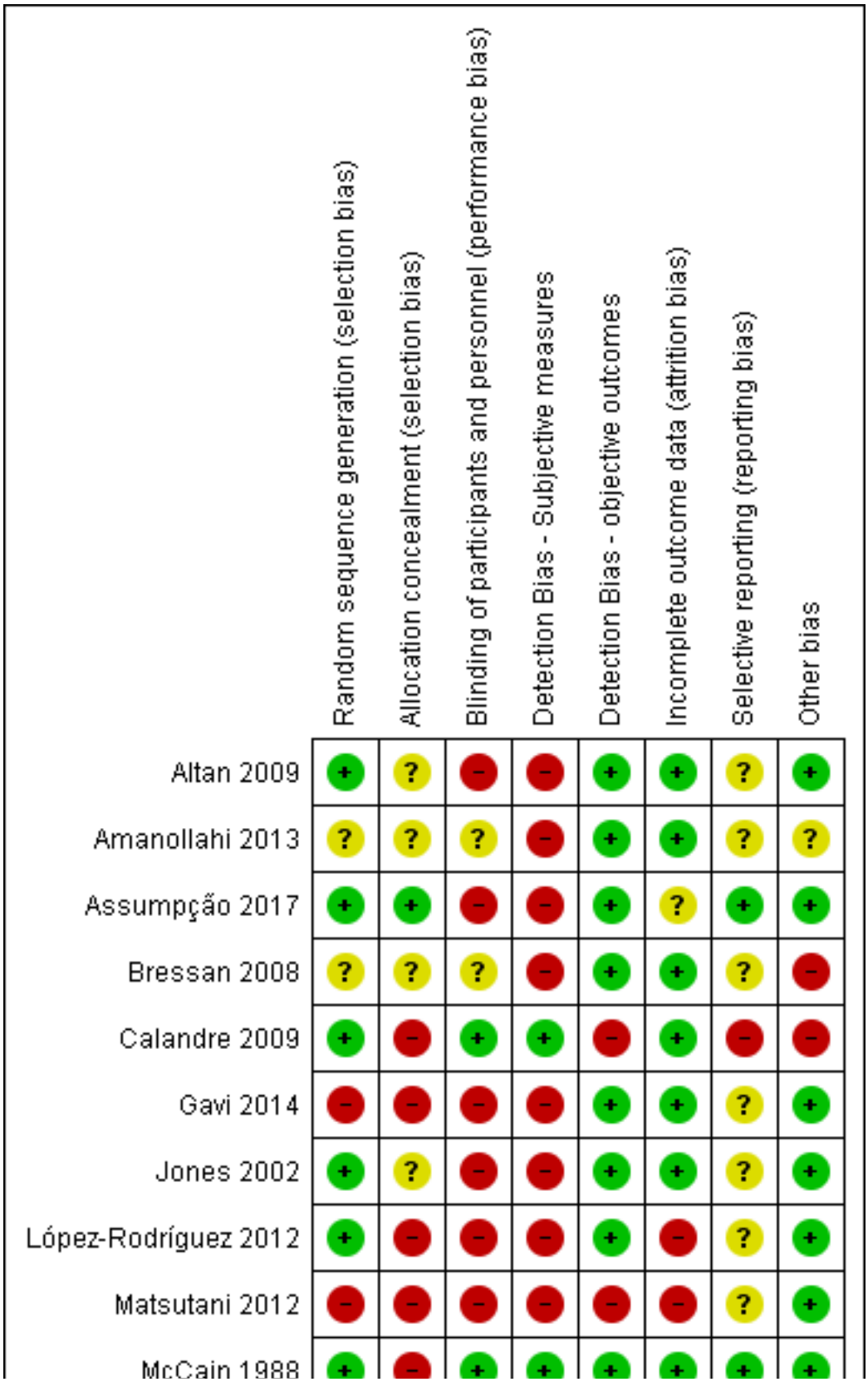
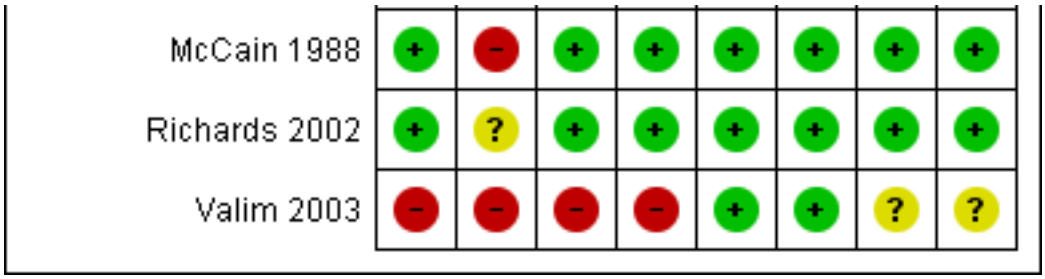


Figure 3. (Continued)



Allocation

Seven of the 12 included studies used an acceptable method of random sequence generation (computer-generated sequence, coin toss, drawing of cards or lots), and were therefore rated as at low risk of bias (Altan 2009; Assumpção 2017; Calandre 2009; Jones 2002; López-Rodríguez 2012; McCain 1988; Richards 2002). In two studies the allocation methods used were unclear (Amanollahi 2013; Bressan 2008). Three studies used unacceptable methods for random sequence generation and were therefore judged to be at high risk of bias (Gavi 2014; Matsutani 2012; Valim 2003). For allocation concealment, we rated one study as at low risk of bias, Assumpção 2017, and five studies as at unclear risk of bias as the information provided was insufficient to permit a definitive judgement. We rated the remaining six studies as at high risk of bias, as allocation was not concealed (i.e. open-label design), or unacceptable methods of allocation concealment (e.g. alternating allocation based on sequence of enrollment or use of a random number list) was employed (Calandre 2009; Gavi 2014; López-Rodríguez 2012; Matsutani 2012; McCain 1988; Valim 2003). Overall, we rated the risk of allocation bias as high (˜50%; Figure 2).

Blinding

In exercise studies, blinding of participants and care providers from treatment allocation is rare.

Performance bias

We rated blinding of participants and personnel (performance bias) as low risk for three studies (Calandre 2009; McCain 1988; Richards 2002); unclear risk for two studies (Amanollahi 2013; Bressan 2008); and high risk for seven studies (Altan 2009; Assumpção 2017; Gavi 2014; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Valim 2003). Overall, we rated risk of performance bias as high (˜55%; Figure 2).

Detection bias

For detection bias, we assessed subjective and objective outcomes separately. Not all trials used a combination of both kinds of outcomes. While completing the 'Risk of bias' tool, we were unable to insert 'not applicable' or to leave the section blank (indicating that the outcome was not measured), thus in such cases we specified 'low risk' and inserted the comment 'not applicable: objective outcomes were not assessed.'

For self-reported outcomes (subjective), we rated all nine studies as at high risk of bias (Altan 2009; Amanollahi 2013; Assumpção 2017; Bressan 2008; Gavi 2014; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Valim 2003). For objectively reported outcomes, four studies blinded outcome assessors to participant group assignment and were therefore rated as at low risk of bias (Altan 2009; Jones 2002; Richards 2002; Valim 2003). We rated six additional studies as at low risk for this domain, however these should actually be rated as 'not applicable' because either the data were not usable or no objective outcomes were measured (Amanollahi 2013; Assumpção 2017; Bressan 2008; Gavi 2014; López-Rodríguez 2012; McCain 1988). Two studies did not blind assessors and were rated as high risk (Calandre 2009; Matsutani 2012). Overall, we rated risk of detection bias as high (100%; Figure 2).

Incomplete outcome data

Nine studies reported complete outcome data. Bressan 2008 had no missing outcome data. Calandre 2009 analyzed data using the intention-to-treat (ITT) principle. Missing outcome data were balanced in numbers across intervention groups, and reasons for missing outcome data were unlikely to be related to true outcomes in Altan 2009, Amanollahi 2013, Gavi 2014, Jones 2002, McCain 1988, and Valim 2003. Richards 2002 replaced missing outcome data with last known value or baseline value. Assumpção 2017 did not use ITT, yet had one participant from the resistance group drop out due to increased pain, thus risk of bias was unclear. López-Rodríguez 2012 and Matsutani 2012 provided incomplete outcome data, therefore we rated these studies as high risk. Overall, we rated risk of attrition bias as low (˜75%; Figure 2).

Selective reporting

Registered protocols were available for four of the included studies (Assumpção 2017; clinicaltrials.gov/ct2/show/NCT01029041; Calandre 2009, clinicaltrials.gov/ct2/show/NCT00550641; Gavi 2014, clinicaltrials.gov/ct2/show/NCT02004405; López-Rodríguez 2012, clinicaltrials.gov/ct2/show/NCT03182556). We rated three studies as having a low risk of reporting bias. One of these three studies had a trial protocol available (Assumpção 2017). Although the remaining two studies rated as at low risk of reporting bias did not have a registered trial protocol, it appeared that published reports included all expected outcomes (McCain 1988; Richards 2002). We rated one study as having a high risk of reporting bias (Calandre 2009). Calandre 2009 had some incongruence between the outcome descriptions and the results reported in the publication. For example, there was no information on tender points as an outcome, yet this was presented in the results. We rated eight out of the 12 included studies as having an unclear risk of bias for this domain (Altan 2009; Amanollahi 2013; Bressan 2008; Gavi 2014; Jones 2002; López-Rodríguez 2012; Matsutani 2012; Valim 2003). Overall, we rated risk of reporting bias as unclear or high (˜75%; Figure 2).

Other potential sources of bias

We rated risk of other potential sources of bias as low (~65%; Figure 2). We assessed eight studies as at low risk of bias for this domain (Altan 2009; Assumpção 2017; Gavi 2014; Jones 2002; López-Rodríguez 2012; Matsutani 2012; McCain 1988; Richards 2002). We rated two studies as at unclear risk of bias due to insufficient information to judge whether an important risk of bias existed (Amanollahi 2013; Valim 2003). We assessed two studies as at high risk of other potential sources of bias. Bressan 2008 had a substantial lack of methodological information to demonstrate rigor in the study design used (e.g. blinding on several levels, allocation, randomization, the instructor/instructors used for the intervention, the level of supervision). Calandre 2009 had baseline imbalances that likely impacted the results.

Effects of interventions

See: **Summary of findings for the main comparison** Flexibility exercise training compared with aerobic exercise training for adults with fibromyalgia

See Summary of findings for the main comparison for the main comparison of flexibility exercise training compared with land-based aerobic exercise training. For comparisons of flexibility exercise training versus untreated controls, resistance exercise training, other comparators and long-term effects of flexibility exercise training and aerobic training, see certainty of evidence in Table 5, Table 6, Table 7, and Table 8.

Flexibility exercise training versus land-based aerobic exercise training at the end of the intervention

Major outcomes

Two studies provided data for HRQoL (Richards 2002; Valim 2003), five studies for pain intensity (Bressan 2008; Matsutani 2012; McCain 1988; Richards 2002; Valim 2003), three studies for fatigue (Bressan 2008; Richards 2002; Valim 2003), one study for physical function (Valim 2003), and one study for the major outcome of stiffness (Bressan 2008). No studies provided clear data for adverse events, and five studies provided data for all-cause withdrawals (Bressan 2008; Matsutani 2012; McCain 1988; Richards 2002; Valim 2003).

Health-related quality of life (self-reported, FIQ Total, lower scores mean better health, negative numbers mean improvement): Two studies provided data for the major outcome HRQoL (Richards 2002; Valim 2003). Assessment of statistical heterogeneity among trials indicated I^2 = 74% (i.e. 50% to 90%: substantial heterogeneity). We evaluated heterogeneity across outcomes for these studies, and since we did not find a large degree of heterogeneity between these studies in other measures we decided to include both studies for the meta-analysis. Mean HRQoL was 46 mm and 42 mm in the flexibility and aerobic groups, respectively. The analysis showed no evidence of a clinically important effect for flexibility exercise training compared with aerobic training postintervention (N = 193; mean difference (MD) 4.14, 95% confidence interval (CI) −5.77 to 14.05; Analysis 1.1). Absolute change was 4% worse (6% better to 14% worse). Relative change in the flexibility groups compared to the aerobic groups was 7.5% worse (10.5% better to 25.5% worse).

Pain intensity (self-reported, 0-to-100 VAS, lower scores mean less pain, negative numbers mean improvement): Data on pain intensity were available for five studies (Bressan 2008; Matsutani 2012; McCain 1988; Richards 2002; Valim 2003). Mean pain was 57 mm and 52 mm in the flexibility and aerobic groups, respectively. The meta-analysis showed no evidence of a clinically important effect with flexibility exercise training compared with aerobic exercise training postintervention (N = 266; MD 4.72, 95% CI −1.39 to 10.83; Analysis 1.2). Absolute change was 5% worse (1% better to 11% worse). Relative change in the flexibility groups compared to the aerobic groups was 6.7% worse (2% better to 15.4% worse). Heterogeneity analysis demonstrated no evidence of heterogeneity (Chi^2 = 0.55, P = 0.55 with df = 2; I^2 = 0%).

Fatigue (self-reported, 0-to-100 scale, lower scores mean less fatigue, negative numbers mean improvement): Three trials assessed fatigue as an outcome (Bressan 2008; Richards 2002; Valim 2003). We did not include data on fatigue provided by Richards 2002 in the meta-analysis as the Chalders fatigue scale was not one of our accepted outcome measures. Mean fatigue was 67 mm and 71 mm in the aerobic and flexibility groups, respectively. The meta-analysis presented no evidence of a clinically important improvement with flexibility exercise training compared to aerobic exercise training postintervention (N = 75; MD −4.12, 95% CI −13.31 to 5.06; Analysis 1.3). Absolute change was 4% better (13% better to 5% worse). Relative change in the flexibility groups compared to the aerobic groups was 6.0% better (19.4% better to 7.4% worse).

Stiffness (self-reported, 0-to-100 FIQ, lower scores mean less stiffness, negative numbers mean improvement): Only one study provided data on stiffness (Bressan 2008). Although the analysis showed a clinically important improvement with flexibility exercise compared with aerobic exercise postintervention (N = 15; MD −29.6, 95% CI −51.47 to −7.73; Analysis 1.4), the 95% confidence interval included both a clinically important and unimportant change. Mean stiffness was 49 mm to 79 mm in the flexibility and aerobic groups, respectively. Absolute change was 30% better (8% better to 51% better). Relative change in the flexibility group compared to the aerobic group was 39% better (10% better to 68% better).

Physical function (self-reported, 0-to-100 FIQ, lower scores means fewer limitations, negative numbers mean improvement): Two studies assessed physical function as an outcome (Bressan 2008; Valim 2003). Data on physical function provided by Bressan 2008 were not presented on a 100-point scale. In addition, there was insufficient information as to how Bressan 2008 reported their data for this particular measure. Consequently, their data were not used for meta-analysis or reported here. Data from Valim 2003 showed no evidence of a clinically important improvement with flexibility exercise compared to aerobic exercise postintervention (N = 60; MD 6.04, 95% CI −3.95 to 16.03; Analysis 1.5). Mean physical function was 23 points and 17 points in the flexibility and aerobic groups, respectively. Absolute change was 6% worse (4% better to 16% worse). Relative change in the flexibility groups compared to the aerobic groups was 14% worse (9.1% better to 37.1% worse).

Adverse events: One adverse effect was described among the 132 participants allocated to flexibility training. The study reported "a patient in the flexibility group had tendinitis of the Achilles tendon, which responded to treatment with local heat and a reduction in exercise for 14 days" (McCain 1988). However, it is unclear whether the tendinitis was related to participation in the intervention.

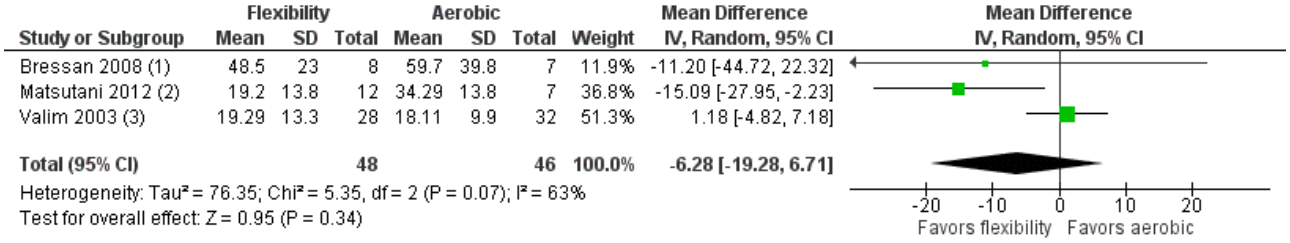
All-cause withdrawal: Rates for flexibility exercise training groups (n1/N1) versus aerobic exercise training groups (n2/N2) were 0/8 versus 0/7 (Bressan 2008) (not included in the analysis); 5/17 versus 8/15 (Matsutani 2012); 2/22 versus 2/20 (McCain 1988); 12/67 versus 12/69 (Richards 2002); and 10/38 versus 6/38 (Valim 2003). We found no evidence of an effect on all-cause withdrawal between the flexibility exercise training and aerobic exercise training groups (risk ratio (RR) 0.97, 95% CI 0.61 to 1.55; Analysis 1.8). Absolute change was 1% fewer withdrawals in the flexibility groups (8% fewer to 21% more). Relative change in the flexibility groups compared to the aerobic groups was 3% fewer (39% fewer to 55% more).

Minor outcomes

Three studies evaluated the effect of flexibility exercise training on the minor outcome of depression (Bressan 2008; Matsutani 2012; Valim 2003), and four studies on tenderness (Matsutani 2012; McCain 1988; Richards 2002; Valim 2003). No studies reported data on improvement in pain greater than 30%.

Depression (self-reported, 0-to-100 FIQ, lower scores mean less depression, negative numbers mean improvement): Data on depression were available for three studies (Bressan 2008; Matsutani 2012; Valim 2003). Assessment of statistical heterogeneity among trials indicated I^2 = 63% (i.e. 50% to 90%: substantial heterogeneity). We investigated the source of this heterogeneity by comparing this meta-analysis to other outcomes in the same comparison. We found no other outcomes indicating substantial heterogeneity, however clinical heterogeneity may be present due to differences in the intervention affecting this outcome, for example length of intervention, frequency of flexibility intervention, type of programs, and sample sizes (i.e. Valim 2003 had a longer intervention of 20 weeks compared to the other studies by Bressan 2008 and Matsutani 2012, which both had interventions of 8 weeks in length; Valim 2003 administered a supervised program 3 times per week, whereas Bressan 2008 and Matsutani 2012 administered a home program with a frequency of 1 time per week; Valim 2003 has a total sample size of 60, whereas Bressan 2008 and Matsutani 2012 had sample sizes of 15 and 19, respectively). The analysis of depression showed absence of an effect postintervention for flexibility exercise training compared with aerobic exercise training (N = 94; MD −6.28, 95% CI −19.28 to 6.71; Figure 4). Relative change in the flexibility groups compared to the aerobic groups was 19.9% better (61% better to 21.2% worse).

Figure 4. Forest plot of comparison: 1 Flexibility vs aerobic (at end of intervention), outcome: 1.6 Depression, 0-63, lower is best (end of intervention).



Footnotes
(1) Flexiblity vs Aerobic. Outcome measured by FIQ (raw data reported in cms. converted to a 100 point scale, lower is best)
(2) Flexiblity vs Aerobic. Outcome measured by Beck Depression Inventory (raw scores on 0 to 63 scale, converted to 100 point scale, lower is best)
(3) Flexiblity vs Aerobic. Outcome measured by Beck Depression Inventory (raw scores on 0 to 63 scale, converted to 100 point scale, lower is best)

Tenderness (0-to-18 TP count, lower score means less tenderness, negative numbers mean improvement): Four trials assessed tenderness. Matsutani 2012, Richards 2002, and Valim 2003 used the tender point count, while McCain 1988 used the total myalgic score. The meta-analysis presented evidence of no effect for flexibility exercise training when compared with aerobic exercise training postintervention (N = 253; standardised mean difference 0.20, 95% CI −0.08 to 0.48; Analysis 1.7). Relative change in the flexibility groups compared to the aerobic groups was 1.4% worse (0.6% better to 3.3% worse).

Improvement in pain greater than 30%: No studies reported data on this outcome.

Flexibility exercise training versus land-based aerobic exercise training, long-term effects

Only one study examined long-term effects (follow-up at 48 weeks, 36 weeks after end of 12-week intervention) and provided data on HRQoL, pain intensity, fatigue, tenderness, and all-cause withdrawals (Richards 2002). Data on stiffness, physical function, and adverse events were not measured at follow-up (Analysis 1.9).

Major outcomes

Health-related quality of life (self-reported, FIQ Total, lower scores mean better health, negative numbers mean improvement): No evidence of an effect was found (N = 135; MD 0.40, 95% CI −5.01 to 5.81).

Pain intensity (self-reported, 0-to-100 VAS, lower scores mean less pain, negative numbers mean improvement): No evidence of an effect was found (N = 136; MD 5.00, 95% CI −2.07 to 12.07).

Fatigue (self-reported, 0-to-100 scale, lower scores mean less fatigue, negative numbers mean improvement): Richards 2002 measured fatigue using the Chadler Fatigue Scale, which was not one of our accepted measures, therefore this information was not included in the review.

Minor outcomes

Tenderness (0-to-18 TP count, lower score means less tenderness, negative numbers mean improvement): We found evidence of an effect between flexibility and aerobic exercise

training favoring aerobic exercise training postintervention (N = 136; MD 2.40, 95% CI 0.66 to 4.14).

Improvement in pain greater than 30%: No studies reported data on this outcome.

Flexibility exercise training versus untreated control at the end of the intervention

Major outcomes

One study provided data for pain intensity, physical function, and all-cause withdrawals (Assumpção 2017). We did not use the data provided for HRQoL, fatigue, and stiffness, which were described as skewed by Assumpção 2017. This study did not provide data on adverse events.

Health-related quality of life (FIQ, SF-36): One study provided data for the major outcome HRQoL (Assumpção 2017), but due to skewing of the data, only medians and interquartile ranges were provided. Although the researchers found within-group improvements in the flexibility group in median total FIQ scores, between-group differences were not statistically significant. The pre-test median scores in the flexibility group of 66.3 points on a 100-point scale dropped to 57.4, versus 73.6 points at pre-test to 72.2 points postintervention in the untreated control group (P = 0.06).

Pain intensity (self-reported, 0-to-100 VAS, lower scores mean less pain, negative numbers mean improvement): One study provided data for the major outcome of pain intensity (Assumpção 2017), and no statistically significant differences between groups were found (N = 28; MD −18.00, 95% CI −37.63 to 1.63; Analysis 2.1). Relative change in the flexibility group compared to the untreated control group was 30% better (2.7% worse to 62.7% better).

Fatigue (FIQ, SF-36): One study provided data for the major outcome of fatigue (Assumpção 2017), but due to skewing of the data, only medians and interquartile ranges were provided. Although the researchers found within-group improvements in the flexibility group in median FIQ fatigue scores, between-group differences were not statistically significant. The pre-test median scores in the flexibility group of 8.6 cm on a 10-centimeter scale dropped to 7.8 cm at post-test, versus 9.2 cm to 8.4 cm in the untreated control group (P = 0.07).

Stiffness (FIQ): Due to skewing of data, one study provided medians and interquartile ranges for stiffness (Assumpção 2017). The pre-test median scores in the flexibility group of 8.3 cm on a 10-centimeter scale dropped to 5.8 cm at post-test, versus 9.2 cm to 9.0 cm in the untreated control group. Between-group differences were not statistically significant.

Physical function (self-reported, 0-to-100 FIQ, lower scores means fewer limitations, negative numbers mean improvement): One study provided data for the major outcome of physical function (Assumpção 2017), and no statistically significant differences between groups were found (N = 28; MD −3.33, 95% CI −16.29 to 9.63; Analysis 2.2). Relative change in the flexibility group compared to the untreated control group was 10.4% better (30.1% worse to 50.9% better).

Adverse events: No adverse event was reported by Assumpção 2017.

All-cause withdrawal: Rates for the flexibility exercise training group (n1/N1) versus the untreated control group (n2/N2) were 4/18 versus 2/16 (Assumpção 2017). We found no significant difference in all-cause withdrawal between flexibility exercise training and the untreated control group (RR 1.78, 95% CI 0.37 to 8.44; Analysis 2.3).

Minor outcomes

One study evaluated the effects of flexibility exercise training on the minor outcome of improvement in pain greater than 30% (Assumpção 2017). We did not use the data for tenderness and depression, which were described as skewed.

Depression (FIQ): Data were not used due to reported skewness.

Tenderness (TP count): Data were not used due to reported skewness.

Improvement in pain greater than 30%: Upon request, Assumpção 2017 provided data for the flexibility exercise training group, but not for the untreated control group.

Flexbility exercise training versus resistance training at the end of the intervention

Major outcomes

One study provided data for HRQoL (Jones 2002), three studies for pain intensity (Assumpção 2017; Gavi 2014; Jones 2002), two studies for fatigue (Gavi 2014; Jones 2002), and two studies for the major outcome of physical function (Gavi 2014; Jones 2002). Three studies provided data on all-cause withdrawals (Assumpção 2017; Gavi 2014; Jones 2002). No study reported complete data on adverse events or measured stiffness.

Health-related quality of life (self-reported, FIQ Total, lower scores mean better health, negative numbers mean improvement): One study provided data for the major outcome HRQoL (Jones 2002); data showed no evidence of an effect of flexibility exercise training compared to resistance training (N = 56; MD 5.55, 95% CI −1.80 to 12.90; Analysis 3.1). Absolute change was 6% worse (2% better to 13% worse). Relative change in the flexibility group compared to the resistance group was 11.5% worse (27.4% worse to 3.8% better).

Pain intensity (self-reported, 0-to-100 VAS, lower scores mean less pain, negative numbers mean improvement): Data on pain intensity were available for three studies (Assumpção 2017; Gavi 2014; Jones 2002). The meta-analysis showed evidence of no effect for flexibility exercise training compared with resistance training (N = 152; MD 1.84, 95% CI −4.15 to 7.83; Analysis 3.2). Absolute change was 2% worse (4% better to 8% worse). Relative change in the flexibility groups compared to the resistance groups was 2.5% worse (11.1% worse to 5.9% better). There was no evidence of heterogeneity for this meta-analysis ($Tau^2 = 0.00$; $Chi^2 = 0.55$, df = 2 (P = 0.76); $I^2 = 0\%$).

Fatigue (self-reported, 0-to-100 scale, lower scores mean less fatigue, negative numbers mean improvement): Two studies assessed fatigue as an outcome (Gavi 2014; Jones 2002). Assessment of statistical heterogeneity among trials indicated $I^2 = 74\%$ (i.e. 50% to 90%: substantial heterogeneity). Some of the clinical heterogeneity may be attributed to differences in the resistance training arm. The meta-analysis showed evidence of no effect for flexibility exercise training versus resistance training

postintervention (N = 122; MD 9.83, 95% CI −5.30 to 24.97; Analysis 3.3). Absolute change was 10% worse (5% better to 25% worse). Relative change in the flexibility groups compared to the resistance groups was 13.1% worse (30.8% worse to 6.54% better).

Physical function (self-reported, 0-to-100 SF-36, converted so that lower scores means fewer limitations, negative numbers mean improvement): Two studies assessed physical function as an outcome (Assumpção 2017; Gavi 2014). Assessment of statistical heterogeneity among studies indicated I^2 = 91% (i.e. 50% to 90%: substantial heterogeneity). Data were checked for accuracy (the SF-36 scale was converted appropriately so that a lower score indicated improvement; the 0-to-30 FIQ scale was converted to a 0-to-100 scale). Given the very large degree of heterogeneity, we did not perform a meta-analysis. Assumpção 2017 compared a 12-week flexibility intervention (N = 14) versus resistance training (N = 16) and found an effect postintervention on physical function favoring the flexibility intervention (FIQ physical functioning; MD −16.66, 95% CI −28.87 to −4.45). Gavi 2014 compared a 16-week flexibility intervention (N = 31) versus resistance training (N = 35) and found an effect postintervention on physical function favoring resistance training (SF-36-Physical capacity; MD 9.47, 95% CI 0.13 to 18.81).

Adverse events: Most studies did not measure adverse events, and other studies reported them incompletely, thus we are uncertain of the estimate. The statement "...arthrosis of the hip" is an adverse event that was reported to have occurred after flexibility exercise training (Gavi 2014), but it is unclear whether the arthrosis was a flare-up related to participation in the intervention.

All-cause withdrawal: Rates for flexibility exercise training groups (n1/N1) versus resistance training groups (n2/N2) were 4/18 versus 3/19 (Assumpção 2017); 9/31 versus 5/35 (Gavi 2014); and 6/28 versus 6/28 (Jones 2002). We found no evidence of effect on all-cause withdrawal between flexibility exercise training and resistance training groups (RR 1.43, 95% CI 0.77 to 2.67; Analysis 3.8).

Minor outcomes

Two studies evaluated the effects of flexibility exercise training on the minor outcome of depression (Gavi 2014; Jones 2002), and one study evaluated the effects on tenderness (Jones 2002).

Depression (self-reported, 0-to-100 FIQ, lower scores mean less depression, negative numbers mean improvement): Data on depression were available for two studies (Gavi 2014; Jones 2002). Data showed no evidence of an effect of flexibility exercise training compared with resistance training postintervention (N = 122; MD 0.47, 95% CI −3.40 to 4.35; Analysis 3.5). Relative change in the flexibility groups compared to the resistance groups was 1.8% worse (16.8% worse to 13.2% better).

Tenderness (0-to-18 TP count, lower score means less tenderness, negative numbers mean improvement): One trial assessed tenderness as an outcome (Jones 2002), showing no evidence of an effect of flexibility exercise training compared to resistance training postintervention (N = 56; MD −0.32, 95% CI −2.03 to 1.39; Analysis 3.6). Relative change in the flexibility group compared to the resistance group was 1.94% better (8.4% worse to 12.3% better).

Improvement in pain greater than 30%: One study evaluated improvement in pain greater than 30% (Assumpção 2017). Rates for the flexibility exercise training group (n1/N1) versus the resistance training group (n2/N2) were 5/14 and 6/16, respectively. We found no evidence of an improvement in pain greater than 30% between the flexibility exercise training and resistance training groups (odds ratio 0.93, 95% CI 0.21 to 4.11; Analysis 3.7).

Flexibility exercise training versus other interventions at the end of the intervention and long term

We did not pool studies as we did not consider interventions to be comparable across trials. Four studies provided data for this comparison (Altan 2009; Amanollahi 2013; Calandre 2009; López-Rodríguez 2012). The comparisons were as follows:

- flexibility exercise training versus Pilates (Altan 2009);
- flexibility exercise training versus Tai Chi (Calandre 2009);
- flexibility exercise training versus aquatic biodanza (López-Rodríguez 2012); and
- flexibility exercise training versus medication (i.e. ibuprofen) and flexibility exercise training versus friction massage (arm 3) (Amanollahi 2013).

Our analyses showed effect sizes on major and minor outcome variables for each of the included studies. Unless otherwise indicated, investigators measured HRQoL, pain, fatigue, and stiffness on a 0-to-100 scale, with lower scores best and negative numbers meaning improvement. Physical function was measured on a 0-to-3 scale, depression on a 0-to-63 scale, and tenderness on a 0-to-18 scale; lower scores are best, and negative numbers mean improvement. No studies reported data on improvement in pain greater than 30%. Four studies provided data on all-cause withdrawals (Altan 2009; Amanollahi 2013; Calandre 2009; López-Rodríguez 2012). Data on adverse events were available from Altan 2009 and Amanollahi 2013, but not always for both study arms.

Flexibility exercise training versus Pilates

End of intervention

Altan 2009 compared a 12-week program of flexibility exercise training (described as "home exercise relaxation and stretching") (n = 25) versus Pilates (n = 25). We found evidence of an effect postintervention favoring Pilates for both HRQoL (FIQ Total, N = 49; MD 14.00, 95% CI 2.50 to 25.50; Analysis 4.1) and pain intensity (VAS; N = 49; MD 19.00, 95% CI 8.28 to 29.72; Analysis 4.2). Altan 2009 found no between-group differences postintervention in tenderness (TP count; N = 49; MD 0.90, 95% CI −1.39 to 3.19) and reported no adverse events (i.e. injuries, exacerbations, or other) in either group. ("We observed no adverse effect of Pilates exercises.") There was no mention of adverse events in the control group (flexibility exercise training and relaxation). All-cause withdrawal rates for the flexibility exercise training group (n1/N1) versus the Pilates group (n2/N2) were 1/24 versus 0/25.

Long term

Altan 2009 provided follow-up data 12 weeks after the end of a 12-week intervention for HRQoL, pain intensity, and tenderness. We found no evidence of a difference between groups for HRQoL (N = 49; MD 8.3, 95% CI −4.84 to 21.4) or tenderness (N = 49; MD 1.1, 95% CI −0.97 to 3.17). However, we found evidence of an effect on pain intensity favoring Pilates (N = 49; MD 13, 95% CI 0.09 to 25.91; Analysis 4.9).

Flexibility exercise training versus Tai Chi

End of intervention

Calandre 2009 compared a 6-week flexibility intervention (in water) (N = 39) versus Tai Chi (in water) (N = 42). We found no evidence of an effect on HRQoL (FIQ Total; N = 81; MD 3.80, 95% CI −2.89 to 10.49; Analysis 4.1); pain intensity (VAS; N = 81; MD 0.00, 95% CI −9.58 to 9.58; Analysis 4.2); fatigue (FIQ VAS; N = 81; MD 3.00, 95% CI −6.83 to 12.83; Analysis 4.3); stiffness (FIQ VAS; N = 81; MD 6.00, 95% CI −5.33 to 17.33; Analysis 4.4); depression (Beck Depression Inventory; N = 81; MD −0.10, 95% CI −2.72 to 2.52; Analysis 4.6); or tenderness (TP count; N = 81; MD −0.50, 95% CI −1.98 to 0.98; Analysis 4.7). Adverse events were not measured for the flexibility group. However, three participants in the Tai Chi group dropped out, two due to “pain exacerbation” and one due to "chlorine hypersensitivity."

Long term

Calandre 2009 provided follow-up data 12 weeks after the end of the 6-week intervention for HRQoL, pain intensity, fatigue, stiffness, depression, and tenderness. We found no evidence of effect between groups in HRQoL (N = 81; MD 2.3, 95% CI −3.69 to 8.29); pain intensity (N = 81; MD −2, 95% CI −11.59 to 7.59); fatigue (N = 81; MD 2, 95% CI −5.62 to 9.62); stiffness (N = 81; MD 0.0, 95% CI −9.37 to 9.37); depression (N = 81; MD −0.31, 95% CI −4.40 to 3.78); and tenderness (N = 81; MD 0.0, 95% CI −1.54 to 1.54). All-cause withdrawal rates for the flexibility exercise training group (n1/N1) versus the Tai Chi group (n2/N2) were 5/39 versus 10/42 (RR 0.54, 95% CI 0.20 to 1.44; Analysis 4.9).

Flexibility exercise training versus aquatic biodanza

End of intervention

López-Rodríguez 2012 compared a 12-week flexibility intervention (N = 20) versus aquatic biodanza (N = 19). We found evidence of an effect favoring aquatic biodanza postintervention on HRQoL (FIQ Total; N = 39; MD 17.07, 95% CI 7.86 to 26.28; Analysis 4.1); fatigue (FIQ VAS; N = 39; MD 11.40, 95% CI 1.09 to 21.71; Analysis 4.3); and stiffness (FIQ VAS; N = 39; MD 14.00, 95% CI 2.68 to 25.32; Analysis 4.4). López-Rodríguez 2012 did not find between-group differences postintervention in physical function (FIQ Activities of Daily Living), 0-to-3-millimeter scale; N = 39; MD 0.37, 95% CI 0.05 to 0.69; Analysis 4.5) or depression (Beck Depression Inventory; N = 39; MD 0.65, 95% CI −3.79 to 5.09; Analysis 4.6). One participant in the flexibility group dropped out of the study due to "worsening of symptom with the training" (information obtained from correspondence with author). No adverse events were reported for the aquatic biodanza group. All-cause withdrawal rates for the flexibility exercise training group (n1/N1) versus the aquatic biodanza group (n2/N2) were 15/35 versus 16/35 (RR 0.94, 95% CI 0.55 to 1.59).

Long term

Long-term effects were not investigated.

Flexibility exercise training versus friction massage

End of intervention

Amanollahi 2013 compared a 4-week flexibility intervention (N = 45) versus friction massage (N = 45). We found evidence of an effect on pain intensity postintervention favoring flexibility (VAS; N = 90; MD −28.00, 95% CI −40.84 to −15.16; Analysis 4.2). Four participants (7%) in the flexibility exercise training group and 11 participants (22.6%) in the friction massage group reported an increase in pain levels. All-cause withdrawal rates for the flexibility exercise training group (n1/N1) versus the friction massage group (n2/N2) were 0/45 versus 0/45 (RR not estimable).

Long term

Long-term effects were not investigated.

Flexibility exercise training versus medication (ibuprofen)

End of intervention

Amanollahi 2013 compared a 4-week flexibility exercise intervention (N = 45) versus medication (ibuprofen) (N = 45). We found no evidence of an effect on pain intensity (VAS; N = 90; MD −8.00, 95% CI −20.21 to 4.21; Analysis 4.2). Five participants in the medication intervention group reported side effects to the ibuprofen medications. All-cause withdrawal rates for the flexibility exercise training group (n1/N1) versus the medication group (n2/N2) were 6/45 versus 0/45 (RR 13.00, 95% CI 0.75 to 224.13).

Long term

Long-term effects were not investigated.

DISCUSSION

This review is one of a series of reviews examining the effects of physical activity interventions for adults with fibromyalgia; this review focused on flexibility exercise training.

Summary of main results

Twelve unique studies involving 743 people met our inclusion criteria. The comparisons were as follows.

- Flexibility exercise training versus untreated controls. One study involving 28 participants compared flexibility exercise training versus control. Results showed no evidence of an effect on pain intensity, physical function, improvement in pain greater than 30%, or all-cause withdrawals. Health-related quality of life, fatigue, and stiffness were not analyzed as data were reported as being skewed. No long-term effects were investigated. The overall certainty of the evidence was low.
- Flexibility exercise training versus land-based aerobic exercise training. Five studies involving a total of 266 participants compared flexibility exercise training versus aerobic exercise training. Although we found evidence of an effect favoring the flexibility exercise group for stiffness (one study), we found no evidence of an effect on HRQoL, pain intensity, fatigue, physical function, all-cause withdrawal, depression, or tenderness. When evaluating long-term effects, we found evidence of an effect of aerobic exercise on tenderness. The overall certainty of the evidence was very low.
- Flexibility exercise training versus resistance training. Three studies involving 152 participants compared flexibility exercise training to resistance training. We found no evidence of an effect for pain intensity, fatigue, depression, all-cause withdrawal, HRQoL, physical function, tenderness, or improvement in pain greater than 30%. Stiffness was not measured. No long-term effects were investigated in any of the studies. The overall certainty of the evidence was low to very low.
- Flexibility exercise training versus other interventions. Four studies involving 299 participants compared flexibility exercise training versus other interventions. Three of these studies had

two parallel arms, and one had three parallel arms. Owing to the differences between interventions and comparators, data were not pooled. In between-group comparisons within single studies comparing flexibility exercise training to a) Pilates, we found evidence of an effect of Pilates on HRQoL and pain intensity, but no evidence of an effect on tenderness; b) Tai Chi, we found no evidence of an effect on HRQoL, pain intensity, fatigue, stiffness, depression, or tenderness; c) aquatic biodanza, we found evidence of an effect of aquatic biodanza on HRQoL, fatigue, and stiffness, but no evidence of an effect on physical function or depression; d) medications, we found no evidence of an effect on pain intensity; and e) friction massage, we found evidence of an effect of flexibility exercise training on pain intensity. These results must be interpreted with caution due to the risk of bias resulting from methodological weaknesses. We assessed the certainty of the evidence for this comparison as very low.

Overall completeness and applicability of evidence

Samples recruited by the included studies consisted mainly of women 35 to 55 years old. Although some men were included, we were unable to calculate a precise number due to lack of information. The 12 included studies were conducted in seven different countries from Europe and North and South America. However, four of the included studies were from Brazil, and the authors of these four studies, Assumpção 2017, Bressan 2008, Matsutani 2012, and Valim 2003 may belong to a joint research group as they are co-authors on each other's studies. Our findings are thus not easily generalizable beyond middle-aged, largely Caucasian (understood to be white), female populations. Sample sizes were small, and pooled samples were still less than the 400 criterion, therefore we recommend caution in generalizing results of this review to the wider population of individuals with fibromyalgia.

Flexibility exercises are often embedded in programs targeting individuals with fibromyalgia within the context of current practice; however, in some instances flexibility exercises may be integrated into the warm-up and/or cool-down regimens rather than being treated as a separate treatment intervention. In our review, some researchers employed flexibility exercises as a control, Altan 2009; López-Rodríguez 2012, or as part of a relaxation intervention, Richards 2002, which may further underscore the lack of recognition of flexibility exercise training as a unique treatment on its own. It is thus plausible that we may have captured only some of the published papers on flexibility exercise and fibromyalgia.

The duration of the flexibility exercise training sessions ranged from 40 to 60 minutes and were a mixture of (unsupervised) home-based programs and supervised group sessions. The flexibility interventions in this review did not meet all recommended FITT (frequency, intensity, time, and type) principles for flexibility exercise training for healthy individuals (see Table 1 and Table 4) (ACSM 2013). Consequently, the benefits of flexibility exercise training may be underestimated in these studies.

According to the 2013 ACSM guidelines for healthy adults (ACSM 2013), the recommended frequency for flexibility training regimens is two to three days per week, with daily being more effective. None of the included studies had a frequency more than three days per week, and ranged from one to three days per week, with frequency fixed throughout the program. Regarding the intensity of the flexibility exercise training program, the 2013 ACSM guidelines recommend the stretch to be taken to the point of tightness or slight discomfort. Eleven of the 12 included studies did not provide information on the intensity of their programs, thus making judgement difficult. The 2013 ACSM guidelines recommend holding the stretch for 10 to 30 seconds. Seven of the 12 included studies met the recommended time for holding each stretch; in four studies this was unclear; and one study did not meet the recommended hold. For type of flexibility exercise, the 2013 ACSM guidelines recommend a series of flexibility exercises for each of the major muscle-tendon units with static, dynamic, ballistic, and proprioceptive neuromuscular facilitation (PNF) all stated as being effective. Most studies met this criteria, with only one study providing insufficient information to permit a judgement. For volume and pattern, the guidelines suggest that a reasonable target is to perform 60 seconds of total stretching for each flexibility exercise with each stretch repeated two to four times. Only three studies met the recommended guidelines, with the remaining studies providing insufficient information to permit a judgement.

Quality of the evidence

The evidence presented in this review was obtained from trials published in academic journals, registered and published RCT protocols, and trial author responses to requests for information. Using the GRADE system of rating evidence for major outcomes, we judged the overall certainty of evidence for the comparison of flexibility exercise training versus the land-aerobic exercise training to be very low after downgrading due to issues related to selection and performance bias, and potential limitations related to inconsistency (i.e. heterogeneity of interventions) or imprecision (i.e. total cumulative sample size lower than 400). The sample sizes of the included trials were often small, and even after pooling the data in the meta-analysis, participant numbers were smaller than desired. In some trials, flexibility was used as a proxy (i.e. flexibility exercise training was used as the control or combined with relaxation), making judgements on the benefits of flexibility exercise training challenging. The available evidence is limited by the number and quality of the included trials, preventing us from reaching robust conclusions regarding the benefits and harms of flexibility exercise training for adults with fibromyalgia. We cannot offer a thorough understanding of adverse effects from flexibility exercise training due to the lack of information provided in the included studies. We found that withdrawal rates did not differ between flexibility and aerobic training. We rated the certainty of the evidence as very low for long-term benefits of flexibility exercise for HRQoL and pain intensity after downgrading for selection bias, indirectness (i.e. flexibility was used along with relaxation as the control), and imprecision (i.e. small number of participants) (see Table 5). For the comparison of flexibility exercise training versus aerobic exercise training, we are thus uncertain whether flexibility exercise training leads to improvements in HRQoL, pain intensity, fatigue, stiffness, and physical function or decreases withdrawals and adverse events because the certainty of the evidence is very low.

For the comparison of flexibility exercise training versus untreated control, there was only one study and the overall certainty of the evidence was low for the measured outcomes (pain intensity and physical function). Selection and performance bias issues as well as imprecision (i.e. total cumulative sample size lower than 400) led to downgrading of the evidence (see Table 6). Withdrawal rates did not

differ between flexibility exercise training and untreated control. Consequently, flexibility exercise training may lead to little or no difference in pain intensity, physical function, and withdrawals.

For the comparison of flexibility exercise training versus resistance training, we found similar issues to the comparison of flexibility exercise training versus aerobic exercise training, which led to downgrading of the evidence for major outcomes (HRQoL, pain intensity, fatigue, and physical function) to low to very low certainty (see Table 7). For this comparison one study reported on the outcome of greater than 30% improvement of pain. The certainty of evidence was low owing to selection and performance bias and small sample size. Flexibility may thus lead to little or no difference in improvement of HRQoL, pain intensity, and pain greater than 30%. We are uncertain whether flexibility improves fatigue and physical function and decreases withdrawals and adverse events because the certainty of the evidence is very low.

For the comparison of flexibility exercise training versus other interventions, the certainty of evidence ranged from low to very low for HRQoL, pain intensity, fatigue, stiffness, and physical function. We downgraded the certainty of the evidence owing to issues related to risk of bias (selection and performance bias), imprecision (small number of participants), and heterogeneity of the interventions (see Table 8). Flexibility may thus lead to little or no difference in physical function, and it is unclear whether flexibility improves HRQoL, pain intensity, fatigue, and stiffness and decreases withdrawals and adverse events because the certainty of the evidence is very low.

Potential biases in the review process

We attempted to control for bias in the review process in the following ways.

- We followed our protocol and documented any deviations from it and reasons for the deviations. We strove for transparency in our decisions and procedures.
- We applied no language restrictions on our search.
- We described inclusion criteria in sufficient detail to avoid inconsistent application in study selection and documented the inclusion criteria. We updated searches periodically and utilized multiple databases.
- By searching clinical trial registries (e.g. ClinicalTrials.gov), we enhanced the opportunity to identify unpublished trials and selective reporting of outcomes. Publication bias may lead to overestimation of treatment effect by up to 12%.
- We contacted primary authors for clarification and additional information where indicated, although responses were not always obtained. We asked our questions in open-ended fashion to avoid leading questions or answers.
- Our team includes multidisciplinary views and range of expertise, which co-create the synthesis of the evidence: our views include library science, systematic reviews and methods, critical appraisal, clinical rheumatology, exercise physiology, physiotherapy, kinesiology, and knowledge translation and lived experience (i.e. consumers).
- We used a standardized procedure to determine selection and inclusion and assessment of studies in the review, and review authors were trained in data extraction.
- Two members of our multidisciplinary team presented the perspective of consumers (i.e. one team member had fibromyalgia and another team member had another rheumatic disease) and brought the perspective of lived experience during the protocol and review process.
- We used intention-to-treat data preferentially.

Agreements and disagreements with other studies or reviews

We found one previous review on flexibility exercise for fibromyalgia (Lorena 2015). The search for the Lorena 2015 review generated five RCTs published between 1986 and 2010. These five studies were assessed for methodological quality using the PEDro scale, which led to one study, Bressan 2008, being excluded for low methodological quality (PEDro scale = 2). Lorena 2015 performed no meta-analysis. One of the four studies included in Lorena 2015 was a thesis at the time of their review (Assumpcao 2010); it has subsequently been published and is included in our review (Assumpção 2017).

All four studies included in Lorena 2015 were included in our review (with the thesis by Assumpcao 2010 being a companion study to Assumpção 2017 (confirmed by thesis author)). Similar to our review, Lorena 2015 observed a greater concentration of studies investigating flexibility in adults with fibromyalgia after the year 2000. We agree with their general conclusions on the flexibility intervention parameters: flexibility training parameters were poorly described with heterogeneity in the time, frequency, and intensity of sessions between studies. We also agree with the statement by Lorena 2015 that there is a "need for further studies to establish the real benefits of the technique, because the majority of published studies shows low methodological quality."

In contrast to our review, Lorena 2015 assessed methodological quality by the PEDro scale (we used the Cochrane 'Risk of bias' tool). Lorena 2015 states that all studies demonstrated improvement in pain intensity, as well as quality of life and physical condition, however our meta-analyses for pain intensity do not support this. In addition, their review only included McCain 1986, which is a preliminary summary for McCain 1988; results from the later study, McCain 1988, were based on a larger sample for the flexibility group. Our study included data from McCain 1988 (we treated McCain 1986 as a companion study in our review). Matsutani 2012 was included in our meta-analysis for studies comparing flexibility and aerobic training, and Jones 2002 and Assumpcao 2010 (companion study of Assumpção 2017) were included in our meta-analysis for studies comparing flexibility and resistance training. Based on our results, the absolute changes and relative improvements show no evidence of an effect for the flexibility groups.

Theadom 2015 conducted a systematic review examining mind-and-body therapy for fibromyalgia. Theadom 2015 categorized their 61 included studies into five broad groups: psychological therapies, biofeedback, mindfulness meditation therapies, movement therapies, and relaxation-based therapies. Two of the 11 included studies within their movement therapies category, which included interventions such as yoga, Tai Chi, and Pilates, were included in our review (Altan 2009; Calandre 2009). In agreement with our review, physical function and pain intensity were used as major outcomes in Theadom 2015. However, fatigue and quality of life were used as their minor outcomes (these were included among our major outcomes). Also in agreement with our review, the authors of Theadom 2015 found very low-certainty evidence for studies investigating the effects of movement therapies, with trial

quality being reduced by unclear details or high risk of allocation concealment and non-blinding of outcome assessors.

There are several interdisciplinary guidelines on the management of fibromyalgia, that are from Europe (EULAR; European League Against Rheumatism) (Macfarlane 2017), Canada (Fitzcharles 2013), Israel (Ablin 2013), and Germany (Arnold 2012; Langhorst 2012; Winkelmann 2012). The most recently revised recommendations are from the EULAR (Macfarlane 2017). Although specific recommendations for flexibility exercise training are not provided, the authors state that the EULAR recommendations are in agreement with recommendations from other countries on the principles of approach to management. They state that there needs to be emphasis on therapy tailored to the individual, and that non-pharmacological therapies should play a first-line of treatment role. As flexibility exercise training is further studied with larger trials in this population, future treatment guidelines may begin to discuss the possible benefits of flexibility training.

Previous Cochrane Reviews of aerobic and resistance training for adults with fibromyalgia identified evidence of an effect associated with exercise training in comparison to controls (Bidonde 2017; Busch 2013). Given that in this review only one study permitted us to evaluate the effects of flexibility training compared to control, these previous reviews could serve as a benchmark which we can use to establish the effects of flexibility. A previous Cochrane Review of aerobic training for fibromyalgia identified evidence of an effect between aerobic training and controls on HRQoL, pain intensity, stiffness, and physical function (Bidonde 2017). If in our review, we found no evidence of an effect between flexibility exercise training and aerobic exercise training for similar outcomes (e.g. HRQoL, pain intensity, and physical function), it may be plausible that flexibility training may also lead to improvement in these same outcomes compared to controls.

AUTHORS' CONCLUSIONS

Implications for practice

One of the main limitations of the trials included in this review may be that the protocols for flexibility exercise training were not set according to American College of Sports Medicine (ACSM) guidelines (ACSM 2013), thus they may not have reached a needed threshold to achieve benefits for participants with fibromyalgia. None of the included studies met all the recommended FITT-VP (frequency, intensity, time, type, volume, and progression) principles for healthy individuals as outlined in the ACSM 2013 guidelines (see Table 4). In particular, intensity was very poorly outlined in the studies, making judgement difficult. In addition, volume and pattern of flexibility exercises were not met in most studies. Furthermore, the intervention modes and muscle groups being stretch were either poorly described or were highly heterogeneous, which may have further underestimated the benefits of flexibility exercise training (see Table 3).

Among the 342 participants allocated to flexibility training, only two adverse effects were described: “symptom worsening due to certain stretching,” López-Rodríguez 2012, and "a patient in the flexibility group had tendinitis of the Achilles tendon, which responded to treatment with local heat and a reduction in exercise for 14 days,” McCain 1988. These adverse effects were poorly described, and information about whether the tendinitis was a pre-existing condition to the intervention was not clear. It was thus difficult to gauge the safety and potential harms of this intervention.

Our results demonstrated no significant long-term benefits of flexibility exercise training. However, only 3 of the 12 included studies investigated long-term effects. The length of the interventions varied widely (6 weeks to 12 weeks), as did the range of follow-up (12 weeks to 36 weeks). Given this variability, it is difficult to comment on the lasting benefits participants may experience with this form of exercise intervention.

Based on our review, we cannot make specific recommendations about the optimal design of flexibility exercise training protocols. A larger body of high-quality studies with clearly outlined flexibility training protocols that meet the recommended FITT principles and ACSM guidelines will help advance our understanding of the benefits and harms of this exercise mode.

Our findings of a lack of significant differences between flexibility exercise and aerobic exercise training and between flexibility exercise and resistance training may be due to the inclusion of stretching in the warm-up and cool-down phases of the aerobic and resistance interventions. Since the inclusion of some stretching in aerobic or resistance training warm-up or cool-down interventions is a standard protocol, excluding for this was not possible. Although the warm-up and cool-down phases of the comparator interventions were relatively brief, their inclusion may have dampened the magnitude of differences found in the statistical comparisons.

Implications for research

Several implications for further research have been derived from this review. We have used the EPICOT approach to describe implications for future researchers as follows (Brown 2006).

Evidence

We found low- to very low-certainty evidence for outcomes comparing flexibility exercise training to aerobic exercise training and resistance training. A common limitation of exercise trials involved participant and personnel blinding. In addition, participant-reported outcome measures were frequently used. The variability of the interventions and the limited number of studies with very few participants in the comparisons between flexibility exercise training and other interventions and controls precluded meta-analysis. Studies with more standardized interventions and detailed reporting of methodologies to minimize bias may improve our evaluation of the certainty of the evidence in the future.

Population

The majority of individuals included in our review were middle-aged women. Based on available information, the population consisted largely of Caucasian (understood to be white) women living in high-income countries, which makes generalizing the results to other contexts difficult.

Severity of disease and level of physical activity prior to the intervention were not clearly defined. In one study, only sedentary women were included, but a definition of sedentary was not provided.

We recommend that researchers identify subgroups within the data so that future meta-analyses can support analysis of effects by age, level of activity, severity of disease, and pain level at baseline.

Intervention

More details with respect to the frequency, intensity, time, volume, and pattern are needed to better judge if prescribed protocols meet the recommendations outlined by the ACSM. Future studies that better document the FITT-VP characteristics will help us understand and compare the true effects of flexibility exercise training protocols. Stretching exercises used in the warm-up and or cool-down phases of exercise interventions used as comparators to flexibility training should also document the details of these components.

The mode of intervention delivery was variable, with some studies having supervised sessions and others performing home-based programs.

Adherence to the exercise protocols is important and can contribute to the efficacy and success of the intervention. Monitoring methods and adherence criteria were poorly documented in the included studies. Future studies that better document these details can further our understanding of possible dose-response relationships between exercise and fibromyalgia symptoms.

Blinding of participants to their group assignment or study hypothesis (or both) is very important, and we recommend that researchers report this information in detail in future trials. This will help increase the robustness of future reviews on this topic.

Comparator

In this review flexibility exercise training was compared to aerobic exercise training, resistance training, other interventions, and controls via direct comparison. The body of evidence would be strengthened by more studies in each category. In particular, our understanding on the benefits and harms of flexibility exercise would be improved if more studies were designed so that flexibility exercise was not treated as part of the 'control' group. Robust studies with flexibility as the main intervention being investigated and compared to controls are needed.

Outcomes

Improved documentation is needed in the area of adverse effects (injuries or absence thereof, the number of events, exacerbations of fibromyalgia symptoms, and other associated adverse effects). This information is critical for clinicians to make informed decisions on the safety and feasibility of interventions and should be reported in a standardized and systematic way.

Few studies investigated the long-term effects of flexibility exercise training, and among the studies that did, the length of follow-up was variable. Future studies should include long-term follow-up using more standardized timelines to allow for meaningful comparison.

In accordance with Cochrane methods and Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) recommendations, we included a 30% improvement in pain intensity for interpreting clinical trial efficacy (Dworkin 2008). Only one study in our review measured this outcome. In keeping with Cochrane recommendations, future studies should ensure this outcome is included.

Timestamp

This review presents data identified up to December 2017; updates will be required as new evidence emerges. This review should be updated in three to five years.

ACKNOWLEDGEMENTS

We would like to acknowledge the following.

- Exercise for Fibromyalgia Cochrane Review Team. The review team contributed to various tasks related to reviews of exercise and fibromyalgia (screening of citations, abstracts, and full-text articles for inclusion/exclusion and related consensus activities; data extraction and related consensus activities; regular attendance at monthly meetings; contributing to discussion regarding relevant concepts; interpretation and discussion of results; translation of knowledge gained; presentation of workshops or seminars when possible).
- Mary Brachaniec, Janet Gunderson, and Anne Lyddiatt (consumer members of the Exercise for Fibromyalgia Cochrane Review Team). Special thanks for contributing to team discussions, helping to identify and select outcome measures, drafting and reviewing the Plain language summary, and providing feedback on the final manuscript.
- Suelen Goes, Julia Bidonde, Emiliana De Omena Bomfim, and Jacobus Christiaan van der Spuy for assisting with translation of studies from Portuguese, Spanish, French, and Farsi, which permitted inclusion screening and data extraction (only some studies used in this review).

REFERENCES

References to studies included in this review

Altan 2009 *{published data only}*

Altan L, Korkmaz N, Bingol U, Gunay B. Effect of pilates training on people with fibromyalgia syndrome: a pilot study. *Archives of Physical Medicine and Rehabilitation* 2009;**90**(12):1983-8.

Amanollahi 2013 *{published data only}*

Amanollahi A, Naghizadeh J, Khatibi A, Hollisaz MT, Shamseddini AR, Saburi A. Comparison of impacts of friction massage, stretching exercises and analgesics on pain relief in primary fibromyalgia syndrome: a randomized clinical trial. *Tehran University Medical Journal* 2013;**70**(10):616-22.

Assumpção 2017 *{published data only}*

* Assumpção A, Matsutani LA, Yuan SL, Santo AS, Sauer J, Mango P, et al. Muscle stretching exercises and resistance training in fibromyalgia: which is better? A three-arm randomized controlled trial. *European Journal of Physical and Rehabilitation Medicine* 2018;**54**(5):663-70. [PUBMED: 29185675]

Assumpção Berssaneti A, Pasqual Marques A. Stretching and Strengthening Muscular for Patients with Fibromyalgia: a Randomized Clinical Trial (Exercicios de Alongamento e Fortalecimento Muscular No Tratamentode Pacientes com Bibromialgia: um Ensaio Clinico Randomizado) [Doctoral Thesis]. Sao Paulo, Brazil: Faculdade de Medicina de Universidade de Sao Paulo, 2010.

NCT01029041. Stretching and strengthening exercise in fibromyalgia patients: a randomized clinical trial. clinicaltrials.gov/show/NCT01029041 (first received 9 December 2009).

Bressan 2008 *{published data only}*

Bressan LR, Matsutani LA, Assumpcao A, Marques AP, Cabral CMN. Effects of muscle stretching and physical conditioning as physical therapy treatment for patients with fibromyalgia. *Revista Brasileira de Fisioterapia [Brazilian Journal of Physical Therapy]* 2008;**12**(2):88-93.

Calandre 2009 *{published data only}*

Calandre EP, Rodriguez-Claro ML, Rico-Villademoros F, Vilchez JS, Hidalgo J, Delgado-Rodriguez A. Effects of pool-based exercise in fibromyalgia symptomatology and sleep quality: a prospective randomised comparison between stretching and Tai Chi. *Clinical and Experimental Rheumatology* 2009;**27**(5 Suppl 56):S21-8.

Gavi 2014 *{published data only}*

* Gavi MB, Vassalo DV, Amaral FT, Macedo DC, Gava PL, Dantas EM, et al. Strengthening exercises improve symptoms and quality of life but do not change autonomic modulation in fibromyalgia: a randomized clinical trial. *PLOS ONE* 2014;**9**(3):e90767.

NCT02004405. Heart rate variability in fibromyalgia—effects of strengthening exercises. clinicaltrials.gov/show/NCT02004405 (first received 9 December 2013).

Jones 2002 *{published data only}*

Jones KD, Burckhardt CS, Clark SR, Bennett RM, Potempa KM. A randomized controlled trial of muscle strengthening versus flexibility training in fibromyalgia. *Journal of Rheumatology* 2002;**29**(5):1041-8.

López-Rodríguez 2012 *{published data only}*

López-Rodríguez MM, Castro-Sánchez AM, Fernández-Martínez M, Matarán-Peñarrocha GA, Rodríguez-Ferrer ME. Comparison between aquatic-biodanza and stretching for improving quality of life and pain in patients with fibromyalgia. *Atencion Primaria* 2012;**44**(11):641-9.

Matsutani 2012 *{published data only}*

Matsutani LA, Assumpcao A, Marques AP. Stretching and aerobic exercises in the treatment of fibromyalgia: pilot study [Exercícios de alongamento muscular e aeróbico notratamento da fibromialgia: estudo piloto]. *Fisioterapia em Movimento* 2012;**25**(2):411-8.

McCain 1988 *{published data only}*

McCain GA. Role of physical fitness training in the fibrositis/fibromyalgia syndrome. *American Journal of Medicine* 1986;**81**(3A):73-7. [PUBMED: 3532784]

* McCain GA, Bell DA, Mai FM, Halliday PD. A controlled study of the effects of a supervised cardiovascular fitness training program on the manifestations of primary fibromyalgia. *Arthritis & Rheumatism* 1988;**31**(9):1135-41.

Richards 2002 *{published data only}*

Richards SC, Scott DL. Prescribed exercise in people with fibromyalgia: parallel group randomised controlled trial. *BMJ* 2002;**325**(7357):185.

Valim 2003 *{published data only}*

Valim V, Natour J, Xiao Y, Pereira AF, Lopes BB, Pollak DF, et al. Effects of physical exercise on serum levels of serotonin and its metabolite in fibromyalgia: a randomized pilot study. *Revista Brasileira de Reumatologia* 2013;**53**(6):538-41.

* Valim V, Oliveira L, Suda A, Silva L, Assis M, Barros Neto T, et al. Aerobic fitness effects in fibromyalgia. *Journal of Rheumatology* 2003;**30**(5):1060-9.

References to studies excluded from this review

Ahlgren 2001 *{published data only}*

Ahlgren C, Waling K, Kadi F, Djupsjobacka M, Thornell L, Sundelin G. Effects on physical performance and pain from three dynamic training programs for women with work-related trapezius myalgia. *Journal of Rehabilitation Medicine* 2001;**33**(4):162-9.

Astin 2003 *{published data only}*

Astin JA, Berman BM, Bausel B, Lee WL, Hochberg M, Forys KL. The efficacy of mindfulness meditation plus Qigong movement therapy in the treatment of fibromyalgia: a randomized controlled trial. *Journal of Rheumatology* 2003;**30**(10):2257-62.

Bailey 1999 *{published data only}*

Bailey A, Starr L, Alderson M, Moreland J. A comparative evaluation of a fibromyalgia rehabilitation program. *Arthritis Care & Research* 1999;**12**(5):336-40.

Bakker 1995 *{published data only}*

Bakker C, Rutten M, van Santen-Hoeufft M, Bolwijn P, van Doorslaer E, van der Linden S. Patient utilities in fibromyalgia and the association with other outcome measures. *Journal of Rheumatology* 1995;**22**:1536-43.

Carson 2010 *{published data only}*

* Carson JW, Carson KM, Jones KD, Bennett RM, Wright CL, Mist SD. A pilot randomized controlled trial of the Yoga of Awareness program in the management of fibromyalgia. Pain 2010; Vol. 151, issue 2:530-9.

Carson JW, Carson KM, Jones KD, Mist SD, Bennett RM. Follow-up of Yoga of Awareness for fibromyalgia: results at 3 months and replication in the wait-list group. *Clinical Journal of Pain* 2012;**28**(9):804-13.

Dawson 2003 *{published data only}*

Dawson KA, Tiidus PM, Pierrynowski M, Crawford JP, Trotter J. Evaluation of a community-based exercise program for diminishing symptoms of fibromyalgia. *Physiotherapy Canada* 2003;**55**:17-22.

Demir-Gocmen 2013 *{published data only}*

Demir-Gocmen D, Altan L, Korkmaz N, Arabaci R. Effect of supervised exercise program including balance exercises on the balance status and clinical signs in patients with fibromyalgia. *Rheumatology International* 2013;**33**(3):743-50.

Field 2003 *{published data only}*

Field T, Delage J, Hernandez Reif M. Movement and massage therapy reduce fibromyalgia pain. *Journal of Bodywork and Movement Therapies* 2003;**7**(1):49-52.

Gandhi 2002 *{published data only}*

Gandhi N, Depauw K, Dolney D, Freson T. Effect of an exercise program on quality of life of women with fibromyalgia. *Women & Therapy* 2002;**25**(2):91-103.

Geel 2002 *{published data only}*

Geel SE, Robergs RA. The effect of graded resistance exercise on fibromyalgia symptoms and muscle bioenergetics: a pilot study. *Arthritis & Rheumatism* 2002;**47**(1):82-6.

Genc 2015 *{published data only}*

Genc A, Tur BS, Aytur YK, Oztuna D, Erdogan MF. Does aerobic exercise affect the hypothalamic-pituitary-adrenal hormonal response in patients with fibromyalgia syndrome?. *Journal of Physical Therapy Science* 2015;**27**(7):2225-31.

Gowans 2002 *{published data only}*

Gowans SE, deHueck A, Abbey SE. Measuring exercise-induced mood changes in fibromyalgia: a comparison of several measures. *Arthritis & Rheumatism* 2002;**47**(6):603-9.

Guarino 2001 *{published data only}*

Guarino P, Peduzzi P, Donta ST, Engel CC, Clauw DJ, Williams DA, et al. A multicenter two by two factorial trial of cognitive behavioral therapy and aerobic exercise for Gulf War veterans' illnesses: design of a Veterans Affairs Cooperative Study (CSP #470). *Controlled Clinical Trials* 2001;**22**:310-32.

Han 1998 *{published data only}*

Han SS. Effects of a self-help program including stretching exercise on symptom reduction in patients with fibromyalgia. *Taehan Kanho* 1998;**37**:78-80.

Hunt 2000 *{published data only}*

Hunt J, Bogg J. An evaluation of the impact of a fibromyalgia self-management programme on patient morbidity and coping. *Advancing in Physiotherapy* 2000;**2**(4):168-75.

Karper 2001 *{published data only}*

Karper WB, Hopewell R, Hodge M. Exercise program effects on women with fibromyalgia syndrome. *Clinical Nurse Specialist* 2001;**15**:67-75.

Kendall 2000 *{published data only}*

Kendall SA, Brolin MK, Soren B, Gerdle B, Henriksson KG. A pilot study of body awareness programs in the treatment of fibromyalgia syndrome. *Arthritis Care & Research* 2000;**13**:304-11.

Kibar 2015 *{published data only}*

Kibar S, Yildiz HE, Ay S, Evcik D, Ergin ES. New approach in fibromyalgia exercise program: a preliminary study regarding the effectiveness of balance training. *Archives of Physical Medicine and Rehabilitation* 2015;**96**(9):1576-82.

Kingsley 2005 *{published data only}*

Kingsley JD, Panton LB, Toole T, Sirithienthad P, Mathis R, McMillan V. The effects of a 12-week strength-training program on strength and functionality in women with fibromyalgia. *Archives of Physical Medicine and Rehabilitation* 2005;**86**(9):1713-21.

Mason 1998 *{published data only}*

Mason LW, Goolkasian P, McCain GA. Evaluation of multimodal treatment program for fibromyalgia. *Journal of Behavioral Medicine* 1998;**21**(2):163-78.

Matsutani 2007 *{published data only}*

Matsutani LA, Marques AP, Ferreira EA, Assumpção A, Lage LV, Casarotto RA, et al. Effectiveness of muscle stretching exercises with and without laser therapy at tender points for patients with fibromyalgia. *Clinical and Experimental Rheumatology* 2007;**25**(3):410-5.

Meiworm 2000 *{published data only}*

Meiworm L, Jakob E, Walker UA, Peter HH, Keul J. Patients with fibromyalgia benefit from aerobic endurance exercise. *Clinical Rheumatology* 2000;**19**(4):253-7.

Mobily 2001 *{published data only}*

Mobily KE, Verburg MD. Aquatic therapy in community-based therapeutic recreation: pain management in a case of fibromyalgia. *Therapeutic Recreation Journal* 2001;**35**:57-69.

Nielens 2000 *{published data only}*

Nielens H, Boisset V, Masquelier E. Fitness and perceived exertion in patients with fibromyalgia syndrome. *Clinical Journal of Pain* 2000;**16**(3):209-13.

Offenbacher 2000 *{published data only}*

Offenbacher M, Stucki G. Physical therapy in the treatment of fibromyalgia. *Scandinavian Journal of Rheumatology* 2000;**Suppl 29**:78.

Oncel 1994 *{published data only}*

Oncel A, Eskiyurt N, Leylabadi M. The results obtained by different therapeutic measures in the treatment of generalized fibromyalgia syndrome. *Tip Fakultesi Mecmuasi* 1994;**57**(4):45-9.

Peters 2002 *{published data only}*

Peters S, Stanley I, Rose M, Kaney S, Salmon P. A randomized controlled trial of group aerobic exercise in primary care patients with persistent, unexplained physical symptoms. *Family Practice* 2002;**19**:665-74.

Pfeiffer 2003 *{published data only}*

Pfeiffer AT, Thompson J, Nelson A, Tucker S, Luedtke C, Finnie S, et al. Effects of a 1.5-day multidisciplinary outpatient treatment program for fibromyalgia: a pilot study. *American Journal of Physical Medicine & Rehabilitation* 2003;**82**(3):186-91.

Piso 2001 *{published data only}*

Piso U, Kuther G, Gutenbrunner C, Gehrke A. Analgesic effects of sauna in fibromyalgia [Analgetische wirkungen der sauna bei der fibromyalgie]. *Physikalische Medizin Rehabilitationsmedizin Kurortmedizin* 2001;**11**(3):94-9.

Rooks 2002 *{published data only}*

Rooks DS, Silverman CB, Kantrowitz FG. The effects of progressive strength training and aerobic exercise on muscle strength and cardiovascular fitness in women with fibromyalgia: a pilot study. *Arthritis & Rheumatism* 2002;**47**:22-8.

Schmidt 2011 *{published data only}*

Schmidt S, Grossman P, Schwarzer B, Jena S, Naumann J, Walach H. Treating fibromyalgia with mindfulness-based stress reduction: results from a 3-armed randomized controlled trial. *Pain* 2011;**152**(2):1838-43.

Thieme 2003 *{published data only}*

Thieme K, Gronica-Ihle E, Flor H. Operant behavioral treatment of fibromyalgia: a controlled study. *Arthritis Care & Research* 2003;**49**:314-20.

Tiidus 1997 *{published data only}*

Tiidus PM. Manual massage and recovery of muscle function following exercise: a literature review. *Journal of Orthopaedic & Sport Physical Therapy* 1997;**25**(2):107-12.

Valencia 2009 *{published data only}*

Valencia M, Alonso B, Alvarez MJ, Barrientos MJ, Ayan C, Sanchez VM. Effects of 2 physiotherapy programs on pain perception, muscular flexibility, and illness impact in women with fibromyalgia: a pilot study. *Journal of Manipulative and Physiological Therapeutics* 2009;**32**(1):84-92.

Vlaeyen 1996 *{published data only}*

Vlaeyen JW, Teeken-Gruben NJ, Goossens ME, Rutten-van Molken MP, Pelt RA, Van Eek H, et al. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I Clinical effects. *Journal of Rheumatology* 1996;**23**(7):1237-45.

Wang 2010 *{published data only}*

Wang C, Schmid CH, Rones R, Kalish R, Yinh J, Goldenberg DL, et al. A randomized trial of tai chi for fibromyalgia. *New England Journal of Medicine* 2010;**363**(8):743-54.

Worrel 2001 *{published data only}*

Worrel LM, Krahn LE, Sletten CD, Pond GR. Treating fibromyalgia with a brief interdisciplinary program: initial outcomes and predictors of response. *Mayo Clinic Proceedings* 2001;**76**(4):384-90.

Additional references

Ablin 2013

Ablin JN, Amital H, Ehrenfeld M, Aloush V, Elkayam O, Langevitz P, et al. Guidelines for the diagnosis and treatment of the fibromyalgia syndrome. *Harefuah* 2013;**152**(12):742-7.

Ablin 2017

Ablin JN, Wolfe F. A comparative evaluation of the 2011 and 2016 criteria for fibromyalgia. *Journal of Rheumatology* 2017;**44**(8):1271-6.

ACSM 2013

American College of Sports Medicine. ACSM's Guidelines for Exercise Testing & Prescription. 9th Edition. Baltimore, MD: Lippincott Williams & Wilkins, 2013.

Ambrose 2015

Ambrose KR, Golightly YM. Physical exercise as non-pharmacological treatment of chronic pain: why and when. *Best Practice & Research: Clinical Rheumatology* 2015;**29**(1):120-30.

Arnold 2008

Arnold LM, Crofford LJ, Mease PJ, Burgess SM, Palmer SC, Abetz L, et al. Patient perspectives on the impact of fibromyalgia. *Patient Education and Counseling* 2008;**73**(1):114-20.

Arnold 2012

Arnold B, Häuser W, Arnold M, Bernateck M, Bernardy K, Brückle W, et al. Multicomponent therapy of fibromyalgia syndrome. Systematic review, meta-analysis and guideline [Multimodale therapie des fibromyalgiesyndroms systematische übersicht, metaanalyse und leitlinie]. *Der Schmerz* 2012;**26**(3):287-90.

Bennett 2009

Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The revised fibromyalgia impact questionnaire (FIQR): validation and psychometric properties. *Arthritis Research & Therapy* 2009;**11**(5):415.

Bennett 2014

Bennett RM, Friend R, Marcus D, Bernstein C, Han BK, Yachoui R, et al. Criteria for the diagnosis of fibromyalgia: validation of the modified 2010 preliminary American College of Rheumatology criteria and the development of alternative criteria. *Arthritis Care & Research* 2014;**66**(9):1364-73. [DOI: 10.1002/acr.22301]

Bergner 1981

Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care* 1981;**19**(8):787-805.

Bernardy 2013

Bernardy K, Klose P, Busch AJ, Choy EH, Hauser W. Cognitive behavioural therapies for fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 9. [DOI: 10.1002/14651858.CD009796.pub2]

Bidonde 2014

Bidonde J, Busch AJ, Webber SC, Schachter CL, Danyliw A, Overend TJ, et al. Aquatic exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2014, Issue 10. [DOI: 10.1002/14651858.CD011336]

Bidonde 2017

Bidonde J, Busch AJ, Schachter CL, Overend TJ, Kim SY, Goes SM, et al. Aerobic exercise training for adults with fibromyalgia. *Cochrane Database of Systematic Reviews* 2017, Issue 6. [DOI: 10.1002/14651858.CD012700]

Boomershine 2012

Boomershine CS. A comprehensive evaluation of standardized assessment tools in the diagnosis of fibromyalgia and in the assessment of fibromyalgia severity. *Pain Research and Treatment* 2012;**2012**:653714.

Brown 2006

Brown P, Brunnhuber K, Chalkidou K, Chalmers I, Clarke M, Fenton M, et al. How to formulate research recommendations. *BMJ* 2006;**333**(7572):804-6. [DOI: 10.1136/bmj.38987.492014.94]

Burckhardt 1991

Burckhardt CS, Clark SR, Bennett RM. The fibromyalgia impact questionnaire: development and validation. *Journal of Rheumatology* 1991;**18**(5):728-33.

Busch 2007

Busch AJ, Barber KAR, Overend TJ, Pelosso PMJ, Schachter CL. Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD012700]

Busch 2013

Busch AJ, Webber SC, Richards RS, Bidonde J, Schachter CL, Schafer LA, et al. Resistance exercise training for fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 12. [DOI: 10.1002/14651858.CD010884]

Busch 2015

Busch AJ, van der Spuy I, Tupper S, Kim SY, Bidonde J, Overend TJ. Whole body vibration exercise for fibromyalgia. *Cochrane Database of Systematic Reviews* 2015, Issue 6. [DOI: 10.1002/14651858.CD011755]

Busija 2011

Busija L, Pausenberger E, Haines TP, Haymes S, Buchbinder R, Osborne RH. Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). *Arthritis Care & Research* 2011;**63 Suppl 11**:S383-412.

Buskila 1990

Buskila D, Gladman DD, Langevitz P, Urowitz S, Smythe HA. Fibromyalgia in human immunodeficiency virus infection. *Journal of Rheumatology* 1990;**17**(9):1202-6.

Caspersen 1985

Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports* 1985;**100**(2):126-131.

Chen 2011

Chen CH, Nosaka K, Chen HL, Lin MJ, Tseng KW, Chen TC. Effects of flexibility training on eccentric exercise-induced muscle damage. *Medicine and Science in Sports and Exercise* 2011;**43**(3):491-500.

Choy 2009a

Choy EH, Mease PJ. Key symptom domains to be assessed in fibromyalgia (outcome measures in rheumatoid arthritis clinical trials). *Rheumatic Disease Clinics of North America* 2009;**35**(2):329-37.

Choy 2009b

Choy EH, Arnold LM, Clauw DJ, Crofford LJ, Glass JM, Simon LS, et al. Content and criterion validity of the preliminary core dataset for clinical trials in fibromyalgia syndrome. *Journal of Rheumatology* 2009;**36**(10):2330-4.

Coelho 2008

Coelho LF. The muscular flexibility training and the range of movement improvement: a critical literature review. *Motricidade* 2008;**4**(4):12.

Costa 2009

Costa PB, Graves BS, Whitehurst M, Jacobs PL. The acute effects of different durations of static stretching on dynamic balance performance. *Journal of Strength and Conditioning Research* 2009;**23**(1):141-7.

Covidence 2015 [Computer program]

Veritas Health Innovation. Covidence. Version accessed 17 June 2017. Melbourne, Australia: Veritas Health Innovation, 2015.

Dadabhoy 2008

Dadabhoy D, Crofford LJ, Spaeth M, Russell IJ, Clauw DJ. Biology and therapy of fibromyalgia. Evidence-based biomarkers for fibromyalgia syndrome. *Arthritis Research & Therapy* 2008;**10**(4):211.

de Weijer 2003

de Weijer VC, Gorniak GC, Shamus E. The effect of static stretch and warm-up exercise on hamstring length over the course of 24 hours. *Journal of Orthopaedic and Sports Physical Therapy* 2003;**33**(12):727-33.

Deare 2013

Deare JC, Zheng Z, Xue CCL, Liu JP, Shang J, Scott SW, et al. Acupuncture for treating fibromyalgia. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.CD007070.pub2]

Decoster 2005

Decoster LC, Cleland J, Altieri C, Russell P. The effects of hamstring stretching on range of motion: a systematic literature review. *Journal of Orthopaedic and Sports Physical Therapy* 2005;**35**(6):377-87.

Deeks 2017

Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors). Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). The Cochrane Collaboration, 2017. Available from www.training.cochrane.org/handbook.

Derry 2016

Derry S, Cording M, Wiffen PJ, Law S, Phillips T, Moore RA. Pregabalin for pain in fibromyalgia in adults. *Cochrane Database of Systematic Reviews* 2016, Issue 9. [DOI: 10.1002/14651858.CD011790.pub2]

Dinerman 1992

Dinerman H, Steere AC. Lyme disease associated with fibromyalgia. *Annals of Internal Medicine* 1992;**117**(4):281-5.

Dworkin 2008

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. *Pain* 2008;**9**(2):105-21.

Feland 2001

Feland JB, Myrer JW, Schulthies SS, Fellingham GW, Measom GW. The effect of duration of stretching of the hamstring muscle group for increasing range of motion in people aged 65 years or older. *Physical Therapy* 2001;**81**(5):1110-7.

Fitzcharles 2013

Fitzcharles MA, Ste-Marie PA, Goldenberg DL, Pereira JX, Abbey S, Choiniere M, et al. 2012 Canadian Guidelines for the diagnosis and management of fibromyalgia syndrome: executive summary. *Pain Research & Management* 2013;**18**(3):119-26. [PUBMED: 23748251]

Garber 2011

Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Medicine and Science in Sports and Exercise* 2011;**43**(7):1334-59.

Ghavidel-Parsa 2015

Ghavidel-Parsa B, Bidari A, Amir Maafi A, Ghalebaghi B. The iceberg nature of fibromyalgia burden: the clinical and economic aspects. *Korean Journal of Pain* 2015;**28**(3):169-76. [PUBMED: 26175876]

Gracely 2003

Gracely RH, Grant MA, Giesecke T. Evoked pain measures in fibromyalgia. *Best Practice & Research. Clinical Rheumatology* 2003;**17**(4):593-609.

GRADEpro GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime). GRADEpro GDT. Version accessed 1 May 2017. Hamiliton (ON): McMaster University (developed by Evidence Prime), 2015.

Guissard 2004

Guissard N, Duchateau J. Effect of static stretch training on neural and mechanical properties of the human plantar-flexor muscles. *Muscle Nerve* 2004;**29**(2):248-55.

Guissard 2006

Guissard N, Duchateau J. Neural aspects of muscle stretching. *Exercise and Sport Sciences Review* 2006;**34**(4):154-8.

Hauser 2012

Hauser W, Wolfe F, Tolle T, Uceyler N, Sommer C. The role of antidepressants in the management of fibromyalgia syndrome: a systematic review and meta-analysis. *CNS Drugs* 2012;**26**(4):297-307. [PUBMED: 22452526]

Hauser 2013

Hauser W, Urrutia G, Tort S, Uceyler N, Walitt B. Serotonin and noradrenaline reuptake inhibitors (SNRIs) for fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2013, Issue 1. [DOI: 10.1002/14651858.CD010292]

Higgins 2011

Hggins JPT, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from www.training.cochrane.org/handbook.

Huskisson 1976

Huskisson EC, Jones J, Scott PJ. Application of visual analogue scale to the measurement of functional capacity. *Rheumatology and Rehabilitation* 1976;**15**(3):185-7.

Huskisson 1983

Huskisson EC, Surrock RD, Tugwell P. Measurement of patient outcome. *British Journal of Rheumatology* 1983;**22**:86-9.

Jadad 1996

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary?. *Control Clinical Trials* 1996;**17**(1):1-12.

Jones 2006

Jones KD, Adams D, Winters-Stone K, Burckhardt CS. A comprehensive review of 46 exercise treatment studies in fibromyalgia (1988-2005). *Health and Quality of Life Outcomes* 2006;**4**:67.

Jones 2015

Jones GT, Atzeni F, Beasley M, Fluss E, Sarzi-Puttini P, Macfarlane GJ. The prevalence of fibromyalgia in the general population: a comparison of the American College of Rheumatology 1990, 2010, and modified 2010 classification criteria. *Arthritis & Rheumatology* 2015;**67**(2):568-75. [PUBMED: 25323744]

Kay 2015

Kay AD, Husbands-Beasley J, Blazevich AJ. Effects of contract-relax, static stretching, and isometric contractions on muscle-tendon mechanics. *Medicine and Science in Sports and Exercise* 2015;**47**(10):2181-90.

Kokkonen 2007

Kokkonen J, Nelson AG, Eldredge C, Winchester JB. Chronic static stretching improves exercise performance. *Medicine and Science in Sports and Exercise* 2007;**39**(10):1825-31.

Langhorst 2012

Langhorst J, Häuser W, Bernardy K, Lucius H, Settan M, Winkelmann A, et al. Complementary and alternative therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline. *Des Schemerz* 2012;**26**(3):311-7.

Lanuez 2011

Lanuez FV, Jacob-Filho W, Lanuez MV, Oliveira AC. Comparative study of the effects of two programs of physical exercises in flexibility and balance of healthy elderly individuals with and without major depression. *Einstein (Sao Paulo)* 2011;**9**(3):307-12.

Lorena 2015

Lorena SB, Lima Mdo C, Ranzolin A, Duarte ÂL. Effects of muscle stretching exercises in the treatment of fibromyalgia: a systematic review. *Sociedade Brasileira de Reumatologia* 2015;**55**(2):167-73.

MacFarlane 1999

MacFarlane GJ, Morris S, Hunt IM, Benjamin S, McBeth J, Papageorgiou AC, et al. Chronic widespread pain in the community: the influence of psychological symptoms and mental disorder on healthcare seeking behavior. *Journal of Rheumatology* 1999;**26**(2):413-9.

Macfarlane 2017

Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Hauser W, Fluss E, et al. EULAR revised recommendations for the management of fibromyalgia. *Annals of the Rheumatic Diseases* 2017;**76**(2):318-28. [PUBMED: 27377815]

McMillian 2006

McMillian DJ, Moore JH, Hatler BS, Taylor DC. Dynamic vs. static-stretching warm up: the effect on power and agility performance. *Journal of Strength and Conditioning Research* 2006;**20**(3):492-9.

Mease 2009

Mease P, Arnold LM, Choy EH, Clauw DJ, Crofford L, Glass JM, et al. Fibromyalgia syndrome module at OMERACT 9: domain construct. *Journal of Rheumatology* 2009;**36**(10):2318–29.

Merskey 1994

Merskey H, Bogduk N. Classification of Chronic pain: Description and Definition of Chronic Pain Syndromes and Definitions of Pain Terms. 2nd Edition. Seattle: IASP Press, 1994.

Mitchell 2012 [Computer program]

Mitchell M. Engauge Digitizer Software. Version 5.1. Available at digitizer.sourceforge.net, 2012.

Moore 2012

Moore RA, Derry S, Aldington D, Cole P, Wiffen PJ. Amitriptyline for neuropathic pain and fibromyalgia in adults. *Cochrane Database of Systematic Reviews* 2012, Issue 12. [DOI: 10.1002/14651858.CD008242.pub2]

Mulholland 2001

Mulholland SJ, Wyss UP. Activities of daily living in non-Western cultures: range of motion requirements for hip and knee joint implants. *International Journal of Rehabilitation Research* 2001;**24**(3):191-8.

Nijs 2013

Nijs J, Roussel N, Van Oosterwijck J, De Kooning M, Ickmans K, Struyf F, et al. Fear of movement and avoidance behaviour toward physical activity in chronic-fatigue syndrome and fibromyalgia: state of the art and implications for clinical practice. *Clinical Rheumatology* 2013;**32**(8):1121-9. [PUBMED: 23639990]

Ormseth 2010

Ormseth MJ, Eyler AE, Hammonds CL, Boomershine CS. Milnacipran for the management of fibromyalgia syndrome. *Journal of Pain Research* 2010;**3**:15-24. [PUBMED: 21197306]

Ost 1987

Ost LG. Applied relaxation: description of a coping technique and review of controlled studies. *Behaviour Research and Therapy* 1987;**25**(5):397-409.

Painter 1999

Painter P, Stewart AL, Carey S. Physical functioning: definitions, measurement, and expectations. *Advances in Renal Replacement Therapy* 1999;**6**(2):110-23.

Peluso 2005

Peluso MA, Guerra de Andrade LH. Physical activity and mental health: the association between exercise and mood. *Clinics (Sao Paulo)* 2005;**60**(1):61-70.

Philadelphia Panel 2001

Philadelphia Panel 2001. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for knee pain. *Physical Therapy* 2001;**81**(10):1675-700.

Queiroz 2013

Queiroz LP. Worldwide epidemiology of fibromyalgia. *Current Pain and Headache Reports* 2013;**17**(356):1-6. [PUBMED: 23801009]

Radford 2006

Radford JA, Burns J, Buchbinder R, Landorf KB, Cook C. Does stretching increase ankle dorsiflexion range of motion? A systematic review. *British Journal of Sports Medicine* 2006;**40**(10):870-5.

Rees 2007

Rees SS, Murphy AJ, Watsford ML, McLachlan KA, Coutts AJ. Effects of proprioceptive neuromuscular facilitation stretching on stiffness and force-producing characteristics of the ankle in active women. *Journal of Strength and Conditioning Research* 2007;**21**(2):572-7.

Reid 2004

Reid DA, McNair PJ. Passive force, angle, and stiffness changes after stretching of hamstring muscles. *Medicine and Science in Sports and Exercise* 2004;**36**(11):1944-8.

RevMan 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rosenthal 1995

Rosenthal R. Writing meta-analytic reviews. *Psychological Bulletin* 1995;**118**:183-92.

Roskell 2011

Roskell NS, Beard SM, Zhao Y, Le TK. A meta-analysis of pain response in the treatment of fibromyalgia. *Pain Practice: the Official Journal of World Institute of Pain* 2011;**11**(6):516-27. [PUBMED: 21199320]

Sharman 2006

Sharman MJ, Cresswell AG, Riek S. Proprioceptive neuromuscular facilitation stretching: mechanisms and clinical implications. *Sports Medicine* 2006;**36**(11):929-39.

Smythe 1979

Smythe HA. Non-articular rheumatism and psychogenic musculoskeletal syndromes. In: McCarty DJ, editor(s). Arthritis and Allied Conditions. Philadelphia: Lea and Febiger, 1979.

Smythe 1981

Smythe HA. Fibrositis and other diffuse musculoskeletal syndromes. Textbook of rheumatology. 1st Edition. Philadelphia: WB Saunders, 1981.

Soriano-Maldonado 2016

Soriano-Maldonado A, Estevez-Lopez F, Segura-Jimenez V, Aparicio VA, Alvarez-Gallardo IC, Herrador-Colmenero M, et al. Association of physical fitness with depression in women with fibromyalgia. *Pain Medicine* 2016;**17**(8):1542-52.

Sterne 2017

Sterne JAC, Egger M, Moher D, Boutron I (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors). Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). The Cochrane Collaboration, 2017. Available from www.training.cochrane.org/handbook.

Theadom 2015

Theadom A, Cropley M, Smith HE, Feigin VL, McPherson K. Mind and body therapy for fibromyalgia. *Cochrane Database of Systematic Reviews* 2015, Issue 4. [DOI: 10.1002/14651858.CD001980.pub3]

Tofferi 2004

Tofferi JK, Jackson JL, O'Malley PG. Treatment of fibromyalgia with cyclobenzaprine: a meta-analysis. *Arthritis & Rheumatism* 2004;**51**(1):9-13. [PUBMED: 14872449]

van Tulder 2003

van Tulder MW, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 2003;**28**(12):1290-9.

Ware 1993

Ware JE, Snow KK, Kosinski M, Gandek G. SF-36 Health Survey Manual and Interpretation Guide. SF-36 Health Survey Manual and Interpretation Guide. Vol. **2**, New York: The Health Institute, New England Medical Center, 1993.

WHO 2017

World Health Organization. Health topics—depression. www.who.int/mental_health/management/depression/en/ 7 April 2017.

Wiffen 2013

Wiffen PJ, Derry S, Moore RA, Aldington D, Cole P, Rice AS, et al. Antiepileptic drugs for neuropathic pain and fibromyalgia—an overview of Cochrane Reviews. *Cochrane*

Database of Systematic Reviews 2013, Issue 11. [DOI: 10.1002/14651858.CD010567.pub2]

Winkelmann 2012

Winkelmann A, Häuser W, Friedel E, Moog-Egan M, Seeger D, Settan M, et al. Physiotherapy and physical therapies for fibromyalgia syndrome. Systematic review, meta-analysis and guideline. *Der Schmerz* 2012;**26**(3):276-86.

Winters 2004

Winters MV, Blake CG, Trost JS, Marcello-Brinker TB, Lowe LM, Garber MB, et al. Passive versus active stretching of hip flexor muscles in subjects with limited hip extension: a randomized clinical trial. *Physical Therapy* 2004;**84**(9):800-7.

Wolfe 1990

Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the Multicenter Criteria Committee. *Arthritis & Rheumatism* 1990;**33**(2):160-72.

Wolfe 1997

Wolfe F, Anderson J, Harkness D, Bennett RM, Caro XJ, Goldenberg DL, et al. Work and disability status of persons with fibromyalgia. *Journal of Rheumatology* 1997;**24**(6):1171-8.

Wolfe 2010

Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. *Arthritis Care & Research* 2010;**62**(5):600-10.

Wolfe 2011

Wolfe F, Hassett AL, Katz RS, Michaud K, Walitt B. Do we need core sets of fibromyalgia domains? The assessment of fibromyalgia (and other rheumatic disorders) in clinical practice. *Journal of Rheumatology* 2011;**38**(6):1104-12.

Wolfe 2016

Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RL, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Seminars in Arthritis and Rheumatism* 2016;**46**(3):319-29. [PUBMED: 27916278]

Woolstenhulme 2006

Woolstenhulme MT, Griffiths CM, Woolstenhulme EM, Parcell AC. Ballistic stretching increases flexibility and acute vertical jump height when combined with basketball activity. *Journal of Strength and Conditioning Research* 2006;**20**(4):799-803.

Yunus 1981

Yunus M, Masi AT, Calabro JJ, Miller KA, Feigenbaum SL. Primary fibromyalgia (fibrositis): clinical study of 50 patients with matched normal controls. *Seminars in Arthritis and Rheumatism* 1981;**11**:151-71.

Zettel-Watson 2011

Zettel-Watson L, Rakovski C, Levin B, Rutledge ON, Jones CJ. Impact of employment and caregiving roles on the well-being of people with fibromyalgia syndrome. *Journal of Musculoskeletal Pain* 2011;**19**:8-17.

References to other published versions of this review

Busch 2001

Busch AJ, Schachter CL, Peloso PM. Fibromyalgia and exercise training: a systematic review of randomized clinical trials. *Physical Therapy Review* 2001;**6**:287-306.

Busch 2002

Busch AJ, Schachter CL, Peloso P, Bombardier C. Exercise for treating fibromyalgia syndrome. *Cochrane Database of Systematic Reviews* 2002, Issue 3. [DOI: 10.1002/14651858.CD003786.pub2]

Busch 2008

Busch AJ, Schachter CL, Overend TJ, Peloso PM, Barber KA. Exercise for fibromyalgia: a systematic review. *Journal of Rheumatology* 2008;**35**(6):1130-44.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Altan 2009

Methods	2 groups: control (FX + relaxation); Pilates Length: 12 weeks; follow-up: 24 weeks Study design: randomized clinical trial with parallel groups
Participants	Female:male: 50:0 Age (years (SD)): 50.0 (8.4); 48.2 (6.5) Inclusion: diagnosis of fibromyalgia (ACR 1990)

Altan 2009 *(Continued)*

	Exclusion: any other rheumatoid disease, unstable hypertension, severe cardiopulmonary problems, or any psychiatric disorders affecting participant compliance Duration of illness (years (SD)): unspecified
Interventions	**Control** (n = 25): home exercise relaxation/stretching; **F**requency: 3/week; **D**uration: 1 h; **I**ntensity: holding each stretch for 6 s and relaxed for 4 s; **M**ode: stretching of cervical, shoulder, thoracic, lumbar, gluteal, leg, and cruris muscle groups **Pilates** (n = 25): **F**requency: 3/week; **D**uration: 1 h (5 min breathing, 10 min warm-up, 35 min conditioning, 10 min cool-down); **I**ntensity: not specified; **M**ode: 9 modules covering postural education, search for neutral position, sitting exercise, "anatalgic exercises," and breathing education. Resistance bands and 26-centimeter Pilates balls were used as supportive equipment. The following components were included in the exercises: resistance and stabilization, flexibility and range of motion, proper body alignment, balance, co-ordination, and body awareness.
Outcomes	Health-related quality of life (FIQ Total), pain intensity (VAS), tenderness (tender point count) Measurements taken at 0, 12, 24 weeks.
Adverse Events	In the control (FX + relaxation) group: not measured or reported In the Pilates group: "no adverse effects of Pilates exercises" p.1987
Adherence	Monitoring methods: monthly monitoring but unreported; adherence criteria: not specified; adherence: attendance rate: attendance rate 96%
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: control: met the criteria for Frequency and Type only
Notes	Country: Turkey Language: English Study author contacted: yes, study author provided details on intervention Trial registry record or protocol available: none found Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"They were assigned randomly into two groups using a random number table by the researcher other than the one who performed the evaluation throughout the study." (page 1984)
Allocation concealment (selection bias)	Unclear risk	Insufficient information on allocation concealment to permit evaluation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants and personnel delivering the intervention were not blinded. Personnel performing the evaluations were blinded. "All participants were asked to give no information to the examiner about their treatment protocol" (page 1984)
Detection Bias - Subjective measures	High risk	Self-report instruments: health related quality of life (FIQ Total), pain intensity (VAS)

Altan 2009 *(Continued)*

Detection Bias - objective outcomes All outcomes	Low risk	Objective measures (TP assessment, strength testing). "Evaluations were performed by the same researcher who was unaware of the groups the participants belonged to." (page 1984)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were balanced in numbers across intervention groups with similar reasons for missing data across groups. "One participant in group two was excluded from the study because she was started on selective serotonin reuptake inhibitor in a psychiatric examination during the study." (page 1985)
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Low risk	Study appears to be free of other sources of bias.

Amanollahi 2013

Methods	3 groups: flexibility (FX); medication; friction massage Length: 4 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 129:0 (analyzed); 129 were included in the final assessment Age (years (SD)): 46.73 (11.33); 46.66 (11.44); 46.65 (12.54) Inclusion: age 40 to 50 years, diagnosis of fibromyalgia (ACR 1990) Exclusion: infection, fever, severe physical impairment, inflammatory diseases, cardiopulmonary diseases, uncontrolled endocrine diseases, allergic diseases, pregnancy, malignancy, and severe psychiatric diseases Duration of illness (years (SD)): unspecified
Interventions	**Flexibility** (n = 45): **F**requency: 3/week; **D**uration: 3 repetitions with 30-second holds; **I**ntensity: not specified; **M**ode: static and non-weight bearing stretching of shoulders blade musculature, paraspinal muscles, neck and low back muscle, hamstrings and calf muscles **Medication** (n = 39): 400 mg ibuprofen (Aria Pharmaceutical Co., Iran) 3 x/day and 25 mg nortriptyline (Darou Pakhsh Pharmaceutical Mfg. Co., Iran) 1/d **Friction massage** (n = 45): 3/week, 3 30-second friction massages using the second and third fingers with a pressure of approximately 0.5 to 1 kg/point on the painful spot so that a mild pallor occurred on the practitioner's nails
Outcomes	Health-related quality of life (life satisfaction), pain intensity (VAS) Measurements taken at 0, 1, 4 weeks.
Adverse Events	Not measured or reported
Adherence	Monitoring methods: not stated; adherence criteria: not stated; adherence: not specified
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: flexibility: met the criteria for Frequency, Time, Type, Volume and Pattern only

Amanollahi 2013 *(Continued)*

Notes	Country: Iran Language: Farsi with English abstract. Methods and results translated into English. Author contact: 6 March 2016 and 27 March 2016 (no author response) Trial registry record or protocol available: none found Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information on the method used to generate the allocation sequence to permit evaluation
Allocation concealment (selection bias)	Unclear risk	Insufficient information on allocation concealment to permit judgement of risk of bias
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	No information in translation or English abstract
Detection Bias - Subjective measures	High risk	Self-report instruments: pain intensity (VAS)
Detection Bias - objective outcomes All outcomes	Low risk	Not applicable: objective outcomes were not assessed
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were balanced in numbers across groups with similar reasons for missing data across groups.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Unclear risk	Insufficient information to assess whether an important risk of bias exists

Assumpção 2017

Methods	3 groups: stretching (FX); resistance (RT); control Length: 12 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 44:0 Age (years (SD)): 47.9 (5.3); 45.7 (67.7); 46.9 (6.5) Inclusion: women aged 30 to 55 referred to fibromyalgia outpatient clinic, diagnosis of fibromyalgia according (ACR 1990) Exclusion: non-controlled systemic disorders (diabetes, hypertension), neurological and musculoskeletal conditions that could compromise assessments, impaired alertness or comprehension, relevant

Assumpção 2017 *(Continued)*

	joint disorders (severe arthritis, arthroplasty of the hip or knee, rheumatoid arthritis), recent changes in physical activity, and recent changes in therapy for fibromyalgia (medication, educational programs, alternative medicine, psychotherapy) Duration of illness (years (SD)): unspecified
Interventions	**Stretching** (n = 14): home exercise relaxation/stretching; **F**requency: 2/week; **D**uration: 40 min; **I**ntensity: early stages 3 reps, from fifth week 4 reps, from ninth week 5 reps; intensity of stretch was gradually increased to point of moderate discomfort and held for 30 s; **M**ode: supervised program focusing on large muscles (triceps surae, gluteus, ischiotibial, paravertebral, latissimus dorsi, hip adductor, pectoralis) **Resistance** (n = 16): **F**requency: 2/week; **D**uration: 40 min (5 min breathing, 10 min warm-up, 35 min conditioning, 10 min cool-down); **I**ntensity: first 2 sessions there was no load; 0.5 kg was added each week if participant identified the effort as slightly intense on the Borg Scale (score = 13); 8 reps; **M**ode: dumbbells for upper limbs and shin pads for lower limbs; exercises targeted triceps surae, quadriceps, hip adductors and abductors, hip flexors, elbow flexors and extensors, pectoralis major, and rhomboids **Control** (n = 14): continued with usual medical treatment
Outcomes	We used the data for which SDs were provided: pain intensity (VAS) and physical function (FIQ-PF; range of scale of 0 to 30 was provided by author through email communication; scale was subsequently converted to 0-to-100 scale for analysis). We did not use the data for major outcomes that were described as skewed: health-related quality of life (FIQ, SF-36), fatigue (FIQ, SF-36), and stiffness (FIQ), nor did we did not use the data for 2 minor outcomes: tenderness (TP count) and depression (FIQ) (see Higgins 2011, Section 7.7.3.5). Measurements taken at 0, 12 weeks.
Adverse Events	Stretching group: not measured or reported Resistance group: authors state "one subject in the resistance group interrupted participant in the study because of worsening pain" (page 12 of downloaded paper)
Adherence	Monitoring methods: not stated; adherence criteria: not specified; adherence: stretching group 22%; resistance 16%; control 13%
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: flexibility: met the criteria for Frequency, Intensity, Time, and Type only
Notes	Country: Brazil Language: English Study author contacted: yes, study author provided details on number of participants with > 30% improvement in pain, range of scale used for physical function on the FIQ, adverse events, and funding Trial registry record or protocol available: clinicaltrials.gov/ct2/show/NCT01029041 Funding source/declaration of interest: not funded

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Drawing of lots "for randomization, each subject drew a paper numbered one, two, or three ... from an urn" (page 7 of downloaded paper)

Assumpção 2017 *(Continued)*

Allocation concealment (selection bias)	Low risk	Because the participants drew a paper from an urn, it is unlikely that the study personnel could predict or control the group allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Protocol states "Open label"; in 'limitations of the study' section authors state "moreover, the evaluator was not blinded in the intervention" (page 12 of downloaded paper).
Detection Bias - Subjective measures	High risk	Self-report instruments: pain intensity (VAS)
Detection Bias - objective outcomes All outcomes	Low risk	Not applicable: tender point data not usable as data were reported as skewed
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	ITT not used. 1 participant from resistance group dropped out due to increased pain.
Selective reporting (reporting bias)	Low risk	The protocol is available and all major and minor outcomes have been reported including any known adverse events.
Other bias	Low risk	The study appears to be free of other sources of bias.

Bressan 2008

Methods	2 groups: stretching (FX); physical conditioning exercises (AE) Length: 8 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 15:0 Age (years (SD)): 49 (7); 44 (8) Inclusion: diagnosis of fibromyalgia (ACR 1990) Exclusion: inadequate cognitive level to understand the orientations and procedures Duration of illness (years (SD)): unspecified
Interventions	**Stretching** (n = 8): **F**requency: 1/week (in addition, stretching at home was recommended); **D**uration: 5 repetitions, with 30 s holds for 40 to 45 min; **I**ntensity: not mentioned; **M**ode: static stretches of triceps surae, ischiotibial, gluteal, paravertebral, latissimocondyloideus, pectoral, trapezius, and respiratory muscles **Physical conditioning exercises** (n = 7): **F**requency: 1/week; **D**uration: 30 min (including 5 min warm-up, 25 min walking, 5 min rest); **I**ntensity: 60% to 75% of the estimated maximum heart rate; **M**ode: walking on motorized treadmill
Outcomes	Pain intensity (FIQ), fatigue (FIQ), stiffness (FIQ), physical function (FIQ), depression (FIQ) Other: feeling well (FIQ), job disability (FIQ), sleep (morning tiredness) (FIQ), anxiety (FIQ) Measurements taken at 0, 8 weeks.
Adverse Events	Not measured or reported

Bressan 2008 *(Continued)*

Adherence	Monitoring methods: heart rate monitoring, but unreported; adherence criteria: not stated; adherence: not stated
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: stretching: met the criteria for Time, Type, Volume, and Pattern only
Notes	Country: Brazil Language: English Study author contacted: no Trial registry record or protocol available: none found Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Insufficient information on the method used to generate the randomization sequence to permit judgement
Allocation concealment (selection bias)	Unclear risk	Insufficient information on allocation concealment to permit judgement of risk of bias
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Insufficient information to permit judgement of risk of bias
Detection Bias - Subjective measures	High risk	Self-report instruments: pain intensity (FIQ), fatigue (FIQ), stiffness (FIQ), physical function (FIQ)
Detection Bias - objective outcomes All outcomes	Low risk	Not applicable: objective outcomes were not assessed
Incomplete outcome data (attrition bias) All outcomes	Low risk	No missing outcome data. No dropouts reported.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	High risk	The level of supervision was not standardized; stretching group was encouraged to do exercises at home but physical conditioning group was not.

Calandre 2009

Methods	2 groups: stretching in water (FX); Tai Chi in water (Tai Chi) Length: 6 weeks; follow-up: 10 weeks and 18 weeks Study design: randomized clinical trial with parallel groups

Calandre 2009 *(Continued)*

Participants	Female:male: 73:8 Age (years (SD)): 51 (8); 49 (8.4) Inclusion: diagnosis of fibromyalgia (ACR 1990) Exclusion: those who had never attended a swimming pool, had disease susceptible to worsen with warm-water exercise such as coronary disease, allergy to chlorine, etc. Duration of illness (years (SD)): 14.1 (8.4); 15.6 (8.7)
Interventions	**Stretching in water** (n = 39): warm shower to condition the body 34.5 °C to 35.5 °C, pool temperature 36 °C; **F**requency: 3 times/week; **D**uration: 60 min (10 min relaxation, 40 min exercise, 10 min relaxation); **I**ntensity: participant selected; **M**ode: active and gentle using 1-meter wooden stick, 1.5-meter flexible tube. Stretches of cervical, upper, and lower extremities and trunk musculature **Tai Chi in water** (n = 42): warm shower to condition the body 34.5 °C to 35.5 °C, pool temperature 36 °C; **F**requency: 3 times/week; **D**uration: 60 min (10 min relaxation, 40 min exercise, 10 min relaxation); **I**ntensity: participant selected; **M**ode: participants were taught the 16 movements that constitute the Tai Chi therapy, using a combination of deep breathing and slow, broad movements of the arms, legs, and torso
Outcomes	Health-related quality of life (HRQoL), pain intensity (VAS), fatigue (FIQ), stiffness (FIQ), physical function (SF-12), tenderness (tender points), depression (BDI) Other: sleep disturbance (Pittsburg Sleep Quality Index), anxiety (State and Trait Anxiety Inventory) Measurements taken at weeks 0, 6, 10, 18 weeks.
Adverse Events	Stretching group:**n**one reported Tai Chi group: pain exacerbation (n = 2); chlorine hypersensitivity (n = 1)
Adherence	Monitoring method: not stated; adherence criteria: not specified adherence: attendance rate 71% (follow-up) 87.1% (completed the program)
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: stretching in water: met the criteria for Frequency and Type only
Notes	Country: Spain Language: English Author contacted: no Trial registry record or protocol available: clinicaltrials.gov/ct2/show/NCT00550641 Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Participants were randomly assigned by means of a computer-generated table of random numbers." (page S14)
Allocation concealment (selection bias)	High risk	"open label design" (page S14)
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Although there was no participant or care provider blinding, the review authors judged that the outcome is not likely to be influenced by lack of blinding, as both groups received an equivalent level of exposure to exercise personnel

Calandre 2009 *(Continued)*

		(Higgins 2011). "As the study was not blinded…" and "A trained physiotherapist, always the same for all of the exercise groups" (page S14)
Detection Bias - Subjective measures	Low risk	Self-report instruments: HRQoL (FIQ total), pain intensity (VAS), fatigue (FIQ), stiffness (FIQ)
Detection Bias - objective outcomes All outcomes	High risk	Objective measures (TP assessment). From author email: "the therapist who performed the tender point assessment was not blinded"
Incomplete outcome data (attrition bias) All outcomes	Low risk	Dropouts specified and ITT used.
Selective reporting (reporting bias)	High risk	Incongruence between outcomes description and results report that likely led to an imbalance in results across groups. No information on TPs as an outcome, yet shows up in the results
Other bias	High risk	Had baseline imbalances: "One of the groups showed significantly better scores of mental health at baseline." (page S18) Statement regarding conflict of interest: "none declared"

Gavi 2014

Methods	2 groups: flexibility (FX); resistance (RT) Length: 16 weeks, follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 66:0 Age (years (SD)): 48.65 (7.6); 44.34 (7.94) Inclusion: women 18 to 65 years of age, diagnosis of fibromyalgia (ACR 1990) Exclusion: cardiovascular, respiratory, metabolic, and rheumatic diseases that could limit exercise, diseases associated with autonomic dysfunction (e.g. arterial hypertension, diabetes, coronary insufficiency), use of medication such as beta blockers, calcium channel blockers, and any other antihypertensive, anticonvulsants, non-tricyclic antidepressants, and opioid analgesics), exercise within last 3 months, inability to understand questionnaires, positive treadmill test, positive treadmill test for myocardial ischemia, receipt of social security benefits Duration of illness (years (SD)): unspecified
Interventions	**Flexibility** (n = 31): **F**requency 2/weeks; **D**uration: 45 min; **I**ntensity: not stated; **M**ode: stretching program included the major muscles **Resistance training** (n = 35): **F**requency: 2/week; **D**uration: 45 min; **I**ntensity: moderate intensity, (overload of 45% of the estimated 1 RM, calculated based on maximal repetitions) in 3 sets of 12 repetitions;**M**ode: 12 dynamic resistance exercises using weight machines for 8 major muscle groups: (quadriceps femoris, hamstrings, biceps brachii, triceps brachii, pectoral, calf, deltoid, and latissimus dorsi) with (leg press, leg extension, hip flexion, pectoral fly, triceps extension, shoulder flexion, leg curl, calf pulldown, shoulder abduction, biceps flexion, and shoulder extension)
Outcomes	Health-related quality of life (FIQ Total), pain intensity (SF-36), fatigue (SF-36), physical function (SF-36), depression (BDI)

Gavi 2014 *(Continued)*

	Other: cardiorespiratory function (max treadmill), cardiorespiratory function submax (6-minute walk test), flexibility (Sit and Reach), physical fitness strength (grip strength, shoulder flexion, leg), mental health (SF-36) Measurements taken at 0, 20 weeks.
Adverse Events	Resistance group: n = 1 lost at follow-up due to "shoulder periarthritis"; however, details as to whether this participant had arthritis prior to the intervention, if it flared up, or if it was aggravated by the intervention was not specified
Adherence	Monitoring method: not stated; adherence criteria: not stated; adherence: 77.5% (follow-up) 100% (completed the program)
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: flexibility: met the criteria for Frequency, Time, and Type only
Notes	Country: Spain Language: English Author contacted: no Trial registry record or protocol available: clinicaltrials.gov/ct2/show/NCT02004405 Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"The patients were sequentially randomized to resistance or flexibility groups according to the order of inclusion in the study" (page 3)
Allocation concealment (selection bias)	High risk	The physician responsible for initial evaluation and inclusion also randomized participants to groups. "The first included patient was allocated to intervention and the second to control group" (page 3)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Only the assessor was blinded to group membership (page 3).
Detection Bias - Subjective measures	High risk	Self-report instruments: HRQoL (FIQ total), pain intensity (SF-36), fatigue (SF-36), physical function (SF-36), depression (BDI)
Detection Bias - objective outcomes All outcomes	Low risk	Not applicable: objective measures were not relevant to this review
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were unlikely to have led to an imbalance in results across groups.
Selective reporting (reporting bias)	Unclear risk	Study protocol was retrospective, i.e. registered after the completion date of study (clinicaltrials.gov/ct2/show/NCT02004405)
Other bias	Low risk	Study appears to be free of other sources of bias.

Jones 2002

Methods	2 groups: stretching (FX); resistance exercises (RT) Length: 12 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 68:0 Age (years (SD)): 46.4 (8.5); 49.2 (6.36) Inclusion: female, 20 to 60 years of age, diagnosis of fibromyalgia (ACR 1990) Exclusion: current or past history of cardiovascular, pulmonary, neurological, endocrine, or renal disease that would preclude involvement in an exercise program, current use of medications such as moderate- or high-dose beta blockers that would significantly affect normal physiological response to exercise, current cigarette smoking, score ≥ 29 on the BDI modified for fibromyalgia, current participation in a regular exercise program Duration of illness (years (SD)): 7.7 (5.5); 6.9 (6.6)
Interventions	**Flexibility** (n = 28): **F**requency: 2/week; **D**uration: 60 min; **I**ntensity: static stretches **M**ode: static stretches for the same 12 major muscle groups as the resistance training group **Resistance training** (n = 28): **F**requency: 2/week; **D**uration: 60 min; **I**ntensity: low, (single sets, initially 4 to 5 repetitions and progressed to 12 repetitions by the end of the study); **M**ode: non-aerobic muscle resistance exercises for 12 muscle groups (gastrocnemius, tibialis anterior, quadriceps, hamstrings, gluteus, abdominals, erector spinae, pectorals, latissimus dorsi and rhomboids, deltoids, biceps and triceps)
Outcomes	Quality of life (QOLS), pain intensity (FIQ-VAS), fatigue (FIQ), stiffness (FIQ), physical functioning (FIQ), depression (BDI and FIQ) Other: morning tiredness (FIQ), anxiety (FIQ), job difficulty (FIQ), and overall well-being in the past week (FIQ), anxiety (Beck Anxiety), muscle strength (maximum isokinetic peak torque of non-dominant side for knee extension flexion, internal and external rotation (Cybex II), body fat was measured in 7 sites (chest, axilla, triceps, sub-scapula, abdomen, supra-iliac, thigh) using a 2-prong spring-loaded caliper (Harpenden) per anthropomorphic standardized guidelines), body weight (kg), self-efficacy (ASES) Measurements taken within 2 weeks before study entry (pre-test) and within 2 weeks after the final exercise class (post-test).
Adverse Events	Not measured or reported
Adherence	Monitoring method: not stated; adherence criteria: not stated; adherence: attendance rate expressed for total study "85% of the participants in both groups (n = 58) attended 13 or more classes" (page 1043)
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: flexibility: met the criteria for Frequency, Time, and Type only
Notes	Country: United States Language: English Author contacted: yes, study author provided information on the sample, intervention, and dropouts Trial registry record or protocol available: none found

Jones 2002 *(Continued)*

	Funding source/declaration of interest: supported by an Individual National Research Service Award (#1F31 NR07337-01A1) from the National Institutes of Health, a doctoral dissertation grant (#2324938) from the Arthritis Foundation, and funds from the Oregon Fibromyalgia Foundation

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomization was accomplished with a coin flip" (page 1042)
Allocation concealment (selection bias)	Unclear risk	Insufficient information on allocation concealment to permit evaluation
Blinding of participants and personnel (performance bias) All outcomes	High risk	Insufficient information, but it is unlikely that participants were blinded. "Both were blinded to group assignment." (page 1042)
Detection Bias - Subjective measures	High risk	Self-report instruments: HRQoL (FIQ total), pain intensity (FIQ), fatigue (FIQ), depression (BDI)
Detection Bias - objective outcomes All outcomes	Low risk	Objective measures: (TP count) "Data were collected by an exercise science technician (strength and body fat) or the principal investigator (all other measures). Both were blinded to group assignment." (page 1042) Other measures: strength (Cybex II isokinetic dynamometer) and body fat (2-prong spring-loaded caliper)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were unlikely to have led to an imbalance in results across groups.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Low risk	There may be risk related to poor adherence to exercise regimen. "85% of the participants attended only slightly more than 50% of the 24 supervised sessions" (Jones 2002, page 1043). The low attendance may have contributed to low power (i.e. type 2 error).

López-Rodríguez 2012

Methods	2 groups: control (FX); biodanza (aerobic exercise in water) (AQ-AE) Length: 12 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: not stated Age (years (SD)): mean age (SD): 55.30 (7.50); 55.50 (7.70) Inclusion: 18 to 65 years of age, diagnosis of fibromyalgia (ACR 1990), willingness to keep their pharmacological treatment constant during the study and not start new exercise or alternative therapies, part of a health center Exclusion: missing 14 or more sessions or changed their pharmacological treatment during the study

López-Rodríguez 2012 *(Continued)*

	Duration of illness (years (SD)): unspecified
Interventions	**Control group** (n = 20): **F**requency: 2/week; **D**uration: 60 min; **I**ntensity: not mentioned; **M**ode: stretching exercises that included global stretches and stretches specific to different muscular areas of the body **Aerobic exercise** (n = 19); **F**requency: 2/week; **D**uration: 60 min (10 min warm-up, 40 min biodanza, 10 min cool-down; **I**ntensity: not mentioned; **M**ode: biodanza in the water with water temperature approximately 29 °C preceded by a shower at 33 °C to 35 °C, biodanza-type movements like walking, slow movements of upper and lower extremities, cool-down stretching
Outcomes	Health-related quality of life (FIQ Total), pain intensity (FIQ), fatigue (FIQ), stiffness (FIQ), physical function (FIQ), tenderness (TP algometry total), depression (BDI) Other: anxiety (FIQ) Measurements taken at weeks 0, 13 weeks.
Adverse Events	Control (flexibility group): n = 1 symptom worsening to certain stretching exercises and dropped from the study; n = 3 worsening of symptoms due to family and work life or season and weather (based on communication from author) Biodanza group: n = 1 ankle fracture; n = 2 worsening of symptoms due to seasonal changes
Adherence	Monitoring method: attendance; adherence criteria: minimum 14 sessions attended; adherence: attendance rate: 57%
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: control group: met the criteria for Frequency and Type only
Notes	Country: Spain Language: Spanish Study author contacted: yes, study author provided information on adverse events, blinding, randomization, intervention, TP evaluation, attrition, and algometry Trial registry record or protocol available: clinicaltrials.gov/ct2/show/NCT03182556 Funding source/declaration of interest: not stated

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	In the translation: "The final sample was made of 70 patients that were randomly assigned to two groups"
Allocation concealment (selection bias)	High risk	No mention in abstract or translation See above; the assignments were "group one and group two"; the participants and researcher in charge of assignment into groups did not know what these groups represented.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Insufficient information, but it is unlikely that participants were blinded

López-Rodríguez 2012 *(Continued)*

Detection Bias - Subjective measures	High risk	Self-report instruments: HRQoL (FIQ total), pain intensity (VAS), fatigue (FIQ), stiffness (FIQ), physical function (FIQ), depression (BDI)
Detection Bias - objective outcomes All outcomes	Low risk	Objective outcomes were not assessed.
Incomplete outcome data (attrition bias) All outcomes	High risk	Missing outcome data likely to have led to an imbalance in results across groups.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Low risk	Study appears to be free of other sources of bias.

Matsutani 2012

Methods	2 groups: stretching (FX); aerobic (AE) Length: 8 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: not stated Age (years (SD)): 49.2 (7.6); 44.1 (7.4) Inclusion: individuals 35 to 60 years of age, able to understand the procedure and follow the basic orientations given, diagnosis of fibromyalgia (ACR 1990) Exclusion: history or suspicion of neoplasia Duration of illness (years (SD)): unspecified
Interventions	**Stretching** (n =17): **F**requency: 1/week with home program of same exercises on same days; **D**uration: 45 min; **I**ntensity: active and gentle. All exercises emphasizing breathing and postural alignment. **M**ode: static stretches held 30 s, repeated 4 times with 30 s rest, progressed from lying to sitting to standing upright or in flexion. Breathing and postural alignment were emphasized. A mirror was used as an aid to the perception of movements of the upper limbs and postural alignment. **Aerobic** (n = 12): **F**requency: daily except the day they had the physical therapy session; **D**uration: 30 min (5 min warm-up, 20 min exercise, 5 min cool-down); **I**ntensity: defined according to heart rate between 60% and 70% for age. **M**ode: treadmill walking and running
Outcomes	Tenderness (TP count), depression (BDI) Others: sleep (Post Sleep Inventory 1 to 13), anxiety-trait (Trait Inventory-State), anxiety-state (Trait Anxiety-State) Measurements taken at 0 and 8 weeks.
Adverse Events	Not measured or reported
Adherence	Monitoring method: attendance; adherence criteria: not specified; adherence: attendance rate: 70.5% (FX) and 46.7% (AE)

Matsutani 2012 *(Continued)*

Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: stretching: met the criteria for Time, Type, Volume, and Pattern only
Notes	Country: Brazil Language: Portuguese Study author contacted: yes, study author provided additional information on exclusion criteria, blinding of participants, exercise interventions, outcome measures, adverse events, and confirmed the lack of a published protocol (see below) Trial registry record or protocol available: none found Funding source/declaration of interest: none stated

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	"The assignment to the groups was made according to the order of arrival, the first patient was randomized and the rest were allocated consecutively in each group. One patient did not want to perform aerobic exercise and therefore was allocated to the stretching group" (page 2 in translated copy)
Allocation concealment (selection bias)	High risk	"Participants were allowed to choose a group if they did not have preference of the exercise group in which they were allocated after randomization. One patient did not want to perform aerobic exercise and therefore was allocated to the stretching group." (page 2 in translated copy)
Blinding of participants and personnel (performance bias) All outcomes	High risk	Neither personnel nor participants were blinded.
Detection Bias - Subjective measures	High risk	Self-report instruments: pain intensity (VAS), depression (BDI)
Detection Bias - objective outcomes All outcomes	High risk	Objective measure: tender point count. In an email for more information, authors stated: "No, neither the assessor nor the patient was blinded during the assessment"
Incomplete outcome data (attrition bias) All outcomes	High risk	Missing outcome data likely led to an imbalance in results across groups.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Low risk	Study appears to be free of other sources of bias.

McCain 1988

Methods	2 groups: flexibility (FX); aerobic exercise (AE) Length: 20 weeks; follow-up: none

McCain 1988 *(Continued)*

Participants	Female:male: mixed, details not stated Age (years SD)): 45.9 (8.2); 35.8 (11.1) Inclusion: diagnosis of fibromyalgia (Smythe 1981), successful treadmill stress test Exclusion: amytriptyline within previous 3 months, ischemic heart disease, symptomatic cardiac arrhythmias, exercise-induced asthma Duration of illness (years (SD)): unspecified
Interventions	**Flexibility** (n = 20): **F**requency: 3/week; **D**uration: 60 min; **I**ntensity: not mentioned; **M**ode: exercises consisted of flexibility maneuvers such that sustained heart rate responses greater than 115 beats per minute were not attained **Aerobic exercise** (n = 18): **F**requency: 3/week; **D**uration: 60 min; **I**ntensity: moderate to vigorous; **M**ode: bicycle ergometry
Outcomes	Pain intensity (VAS), physical function (submaximal cardiorespiratory fitness cycle ergometry), tenderness (TMS) Others: sleep (sleep diary), psychological function (SCL-90-R), Global (Global Severity Index) Measurements taken at weeks 0 and 20 weeks.
Adverse Events	Flexibility group: "a patient in the flexibility group had tendinitis of the achilles tendon, which responded to treatment with local heat and a reduction in exercise for 14 days" (page 1138 but it is unclear whether the tendinitis was related to intervention participation) Aerobic group: none reported
Adherence	Monitoring method: attendance; adherence criteria: attendance rate; attendance rate: 90%
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: flexibility: met the criteria for Frequency only
Notes	Country: Canada Language: English Study author contacted: no Trial registry record or protocol available: none found Funding source/declaration of interest: none reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Consecutive patients were given an odd or even number from a list of random and 'randomized' to receive higher cardiovascular training (even numbers) or flexibility exercise training (odd numbers) for a period of 20 weeks" (page 1136)
Allocation concealment (selection bias)	High risk	A list of random numbers was used (page 1136).

McCain 1988 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	Low risk	"No contact between members of the 2 groups was allowed during the treatment period; patients in one group were unaware of the kind of exercises offered to patients in the other group" (page 1136)
Detection Bias - Subjective measures	Low risk	Self-report instruments: pain intensity (VAS)
Detection Bias - objective outcomes All outcomes	Low risk	Objective outcomes were not assessed.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Reasons for missing outcomes are unlikely to be related to true outcomes. Missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups.
Selective reporting (reporting bias)	Low risk	Study protocol is available, and most of the study's prespecified outcomes have been reported, except variances on the fatigue scale SCL-90-R (page 1139).
Other bias	Low risk	There were some baseline imbalances regarding age, pain intensity scores on VAS (higher in CVR group), sex (both, and the only 2 men in the flexibility group dropped out, also there was a sex difference between groups after the dropouts), however we feel that overall there was a low risk of bias.

Richards 2002

Methods	2 groups: relaxation and flexibility (FX); aerobic exercise (AE) Length: 12 weeks; follow-up: 24 and 48 weeks Study design: randomized clinical trial with parallel groups
Participants	Female:male: not stated Age (years (SD)): 45 (median); 48 (median) Inclusion: men and women aged 18 to 70 years of age, diagnosis of fibromyalgia (ACR 1990), able to to give informed consent Exclusion: individuals for whom an alternative medical diagnosis could explain current symptoms, inability to attend classes, severe pulmonary, cardiovascular, renal, neurological disease precluding involvement in aerobic exercise, inability to co-operate Duration of illness (years (range)): 4 (3 to 4); 6.5 (3 to 10)
Interventions	**Relaxation and flexibility** (n = 67; control treatment): **F**requency: 2/week; **D**uration: 60 min; **M**ode: static and comprised of upper and lower limb stretches and relaxation techniques based on the published regimen by Ost 1987. As the classes continued, more techniques were introduced progressing through progressive muscle relaxation, release-only relaxation and visualization, cue-controlled relaxation, and differential relaxation **Aerobic exercise** (n = 69): **F**requency: 2/week; **D**uration: progressed to 60 min; **I**ntensity: at an intensity that made them sweat slightly while being able to talk comfortably in complete sentences. **M**ode: individualized treadmill walking or cycle ergometry Measurements taken at 0, 12, 24, 48 weeks.

Richards 2002 *(Continued)*

Outcomes	Health-related quality of life (SF-36), pain intensity (McGill Pain Questionnaire, VAS), fatigue (Chalders Fatigue Scale), tenderness (TP count)
Adverse Events	None reported.
Adherence	Monitoring method: attendance; adherence criteria: not specified; attendance rate: "53% of participants attended over one third of the classes" (page 2)
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: relaxation and flexibility: met the criteria for Frequency and Type only
Notes	Country: United Kingdom Language: English Study author contacted: yes, authors provided information on data outcome measures, means and standard deviations Trial registry record or protocol available: none found Funding source/declaration of interest: this study was funded by a Research Training Fellowship of the London region of the NHS executive; no conflict reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"an independent researcher not involved in the assessment used a random number table for allocation"
Allocation concealment (selection bias)	Unclear risk	Insufficient information on allocation concealment to permit evaluation of risk of bias
Blinding of participants and personnel (performance bias) All outcomes	Low risk	The interventions were carried out by personal trainers blinded to the hypothesis of the trial. "They received standardized advise including an explanation of fibromyalgia and encouragement and were told that the exercise offered through prescription would improve their condition"
Detection Bias - Subjective measures	Low risk	Self-report instruments: HRQoL (SF-36), pain (McGill Pain Questionnaire, VAS)
Detection Bias - objective outcomes All outcomes	Low risk	Objective measures: TP assessment: "To give a tender point count a blinded observer recorded tenderness at the 18 sites specified in the fibromyalgia classification criteria." "a single blinded assessor (SR) who recorded physical outcome measures and remained unaware of allocation throughout the trial" (page 2)
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were replaced with last known value or baseline value. 12 participants in each group dropped out; reasons for dropout were not provided.
Selective reporting (reporting bias)	Low risk	Not all results were published, but we obtained all unpublished results we had requested.
Other bias	Low risk	Study appears to be free of other sources of bias.

Valim 2003

Methods	2 groups: stretching exercise (FX); aerobic exercise (AE) Length: 20 weeks; follow-up: none Study design: randomized clinical trial with parallel groups
Participants	Female:male: 76:0 Age (years (SD)): 44 (11); 47 (10) Inclusion: sedentary women, 18 to 60 years of age, diagnosis of fibromyalgia (ACR 1990), never previously treated, newly diagnosed Exclusion: cardiorespiratory disorders limiting exercise, neurological disorders, body mass index > 35, hyperthyroidism, other rheumatologic diseases Duration of illness (years (SD)): unspecified
Interventions	**Stretching exercise** (n = 28): **F**requency: 3/week; **D**uration: 45 min; **I**ntensity: 30 s/stretch for each maximum position; **M**ode: stretching program included 17 exercises using both muscles and joints in a general way, including face, cervical, trunk, and extremities **Aerobic exercise** (n = 32): **F**requency: 3/week; **D**uration: 45 min (5 to 10 min warm-up, 5 min cool-down; **I**ntensity: moderate to vigorous; **M**ode: walking
Outcomes	Health-related quality of life (FIQ Total), pain intensity (VAS), physical function (SF-36), cardiorespiratory maximal and submaximal treadmill test, sit-and-reach test, tenderness (TP count), depression (BDI) Others: anxiety (State-Trait Anxiety Inventory) Measurements taken at 0 and 20 weeks.
Adverse Events	None reported.
Adherence	Monitoring method: attendance; adherence criteria: not specified; adherence: 78.9%
Congruence of EX protocol with ACSM criteria for flexibility	ACSM 2013: stretching exercise: met the criteria for Frequency, Time, and Type only
Notes	Country: Brazil Language: English Study author contacted: yes, clarification received September 2010 by email regarding outcomes Trial registry record or protocol available: none found Funding sources/declaration of interest: FAPESP (State of Sao Paulo) funding, no conflicts reported

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Selection was based on date of admission (page 539 of **2013** paper).
Allocation concealment (selection bias)	High risk	Alternation methods were used to allocate participants (page 1061 of **2003** paper; page 539 of **2013** paper).

Valim 2003 *(Continued)*

Blinding of participants and personnel (performance bias) All outcomes	High risk	Based on correspondence with author, participants were blinded to the hypothesis, and contact with care provider of other interventions was restricted during the study. It is likely personnel were aware of intervention.
Detection Bias - Subjective measures	High risk	Self-report instruments: HRQoL (FIQ total), pain intensity (VAS), physical function (SF-36)
Detection Bias - objective outcomes All outcomes	Low risk	Objective measures: TP assessment: "patients were evaluated by a blinded investigator at the beginning and after 10 and 20 weeks (end of exercise program) in relation to the improvement of aerobic fitness, flexibility, pain, function, quality of life, depression, and anxiety levels" (page 1061) Other objective measures: health professional-rated disease severity/change, cardiorespiratory fitness, flexibility testing
Incomplete outcome data (attrition bias) All outcomes	Low risk	Missing outcome data were balanced in numbers across intervention groups, with similar reasons for missing data across groups.
Selective reporting (reporting bias)	Unclear risk	Insufficient information to permit judgement of risk of bias
Other bias	Unclear risk	Insufficient information to assess whether an important risk of bias exists

ACR: American College of Rheumatology; ACSM: American College of Sports Medicine; AE: aerobic exercise; AQ: aquatic; ASES: Arthitis Self Efficacy Scale; BDI: Beck Depression Inventory; CVR: cardiovascular; FIQ: Fibromyalgia Impact Questionnaire; FIQ-PF: Fibromyalgia Impact Questionnaire: Physical Function; FX: flexibility exercise; HRQoL: health-related quality of life; ITT: intention-to-treat; QOLS: Quality of Life Scale; RT: resistance training; SD: standard deviation; SF-12: 12-item Short Form Health Survey; SF-36: 36-item Short Form Health Survey; SCL-90-R: Symptom Checklist-90-Revised; TMS: Total Myalgic Score; TP: tender point; VAS: Visual Analogue Scale

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Ahlgren 2001	Diagnosis—trapezius myalgia
Astin 2003	Did not meet exercise criteria (QiGong)
Bailey 1999	1-group design
Bakker 1995	Between-group analysis not done
Carson 2010	Intervention did not include flexibility according to our review criteria.
Dawson 2003	1-group before-after design
Demir-Gocmen 2013	Effects of flexibility cannot be isolated.
Field 2003	Intervention included a massage component, thus deemed a more complex intervention.
Gandhi 2002	Not randomized—3-group design: (1) non-exercising control (n = 12), (2) hospital-based exercise group (n = 10), (3) home-based videotaped exercise program (n = 10)
Geel 2002	Not randomized

Study	Reason for exclusion
Genc 2015	Effects of flexibility cannot be isolated.
Gowans 2002	Focuses on measurement issues of selected variables already reported in an included study; new variables did not include standard deviations.
Guarino 2001	Diagnosis—Gulf War Syndrome
Han 1998	Not randomized (geographic control)
Hunt 2000	Diagnosis of fibromyalgia was not clear, even when the author was contacted to clarify the diagnostic criteria that were used.
Karper 2001	Not randomized (program evaluation)
Kendall 2000	Did not meet flexibility exercise criteria (Body Awareness)
Kibar 2015	Effects of flexibility cannot be isolated.
Kingsley 2005	Diagnosis of fibromyalgia made by physician or rheumatologist, but when contacted the authors did not verify the use of published criteria (e.g. ACR 1990 classification).
Mason 1998	Not randomized (participants enrolled in a multimodal treatment compared to participants who were unable to participate due to insurance reasons)
Matsutani 2007	Both groups did stretching, therefore the effects of stretching cannot be isolated.
Meiworm 2000	Not randomized (participants self-selected their group)
Mobily 2001	Case study
Nielens 2000	Not randomized (cross-sectional case control study of fitness)
Offenbacher 2000	Non-experimental—narrative review
Oncel 1994	Insufficient description of exercise (1 group received "medical therapy and exercise"; no further information about the exercise intervention given)
Peters 2002	Not a published diagnosis—"Persistent unexplained symptoms"
Pfeiffer 2003	1-group before-after design
Piso 2001	Not randomized—our translator reported: "The authors wrote only how they recruited nine of the patients. They wrote nothing about if and how the patients were allocated to the two groups." Several attempts to contact the authors for clarification were unsuccessful.
Rooks 2002	1-group design
Schmidt 2011	Intervention did not include flexibility according to our review criteria for intervention type.
Thieme 2003	Did not meet exercise criteria (passive physical therapy with light movement in water—the active exercise was too small a component, not described or quantified sufficiently)
Tiidus 1997	1-group repeated-measures design
Valencia 2009	Effects of flexibility cannot be isolated.

Study	Reason for exclusion
Vlaeyen 1996	Insufficient description of the mode of exercise: "Each session ended with a physical exercise such as swimming or bicycling, excluding systematic physical or fitness training."
Wang 2010	Flexibiity group had an extensive educational component, thus effects of flexibility cannot be isolated.
Worrel 2001	1-group design

ACR: American College of Rheumatology

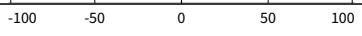
DATA AND ANALYSES

Comparison 1. Flexibility versus aerobic (end of intervention)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention)	2	193	Mean Difference (IV, Random, 95% CI)	4.14 [-5.77, 14.05]
2 Pain, Intensity, 0-100, lower is best (end of intervention)	4	131	Mean Difference (IV, Random, 95% CI)	2.78 [-6.29, 11.85]
3 Fatigue, 0-100, lower is best (end of intervention)	2	75	Mean Difference (IV, Random, 95% CI)	-4.12 [-13.31, 5.06]
4 Stiffness, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
5 Physical function, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
6 Depression, 0-100, lower is best (end of intervention)	3	94	Mean Difference (IV, Random, 95% CI)	-6.28 [-19.28, 6.71]
7 Tenderness 0-18, lower is best (end of intervention)	4	253	Std. Mean Difference (IV, Random, 95% CI)	0.20 [-0.08, 0.48]
8 Withdrawals	5	301	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.61, 1.55]
9 Long-term effects	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
9.1 HRQoL	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.2 Pain	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.3 Tenderness	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 1.1. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention).

Study or subgroup	Favors flexibility N	Mean(SD)	Favors aerobic N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Richards 2002	65	54.7 (14.6)	68	55 (15.2)		56.45%	-0.3[-5.36,4.76]
Valim 2003	28	40.3 (15.5)	32	30.4 (19.2)		43.55%	9.9[1.11,18.69]
Total ***	**93**		**100**			**100%**	**4.14[-5.77,14.05]**
Heterogeneity: Tau²=38.63; Chi²=3.89, df=1(P=0.05); I²=74.26%							
Test for overall effect: Z=0.82(P=0.41)							



Analysis 1.2. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Mean(SD)	Aerobic N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Bressan 2008	8	55.6 (22.6)	6	60.3 (19.2)		13.67%	-4.7[-26.64,17.24]
Matsutani 2012	12	51 (27)	7	63 (17)		16.05%	-12[-31.8,7.8]
McCain 1988	20	49.6 (15)	18	45.4 (15)		38.78%	4.2[-5.35,13.75]
Valim 2003	28	46 (21.8)	32	34.2 (25)		31.51%	11.8[-0.04,23.64]
Total ***	**68**		**63**			**100%**	**2.78[-6.29,11.85]**
Heterogeneity: Tau²=31.5; Chi²=4.78, df=3(P=0.19); I²=37.25%							
Test for overall effect: Z=0.6(P=0.55)							

Analysis 1.3. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention).

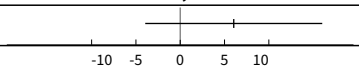
Study or subgroup	Flexibility N	Mean(SD)	Aerobic N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Bressan 2008	8	71.5 (16.7)	7	82.9 (17.5)		27.95%	-11.4[-28.78,5.98]
Valim 2003	28	57.7 (22.3)	32	59 (20.2)		72.05%	-1.3[-12.12,9.52]
Total ***	**36**		**39**			**100%**	**-4.12[-13.31,5.06]**
Heterogeneity: Tau²=0; Chi²=0.93, df=1(P=0.33); I²=0%							
Test for overall effect: Z=0.88(P=0.38)							

Analysis 1.4. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 4 Stiffness, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Mean(SD)	Aerobic N	Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Bressan 2008	8	49.5 (20.3)	7	79.1 (22.6)		0%	-29.6[-51.47,-7.73]

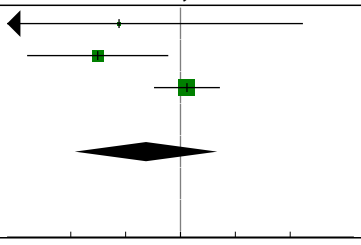
Analysis 1.5. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 5 Physical function, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Aerobic N	Aerobic Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Valim 2003	28	31.7 (21.5)	32	25.7 (17.4)		6.04[-3.95,16.03]



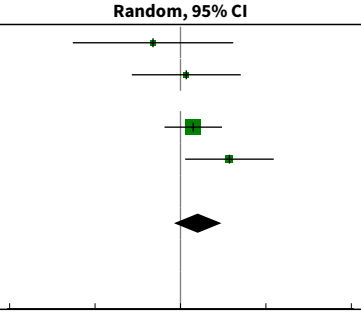
Analysis 1.6. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 6 Depression, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Aerobic N	Aerobic Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Bressan 2008	8	48.5 (23)	7	59.7 (39.8)		11.92%	-11.2[-44.72,22.32]
Matsutani 2012	12	19.2 (13.8)	7	34.3 (13.8)		36.81%	-15.09[-27.95,-2.23]
Valim 2003	28	19.3 (13.3)	32	18.1 (9.9)		51.27%	1.18[-4.82,7.18]
Total ***	**48**		**46**			**100%**	**-6.28[-19.28,6.71]**
Heterogeneity: Tau^2=76.35; Chi^2=5.35, df=2(P=0.07); I^2=62.58%							
Test for overall effect: Z=0.95(P=0.34)							



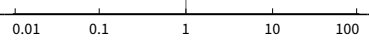
Analysis 1.7. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 7 Tenderness 0-18, lower is best (end of intervention).

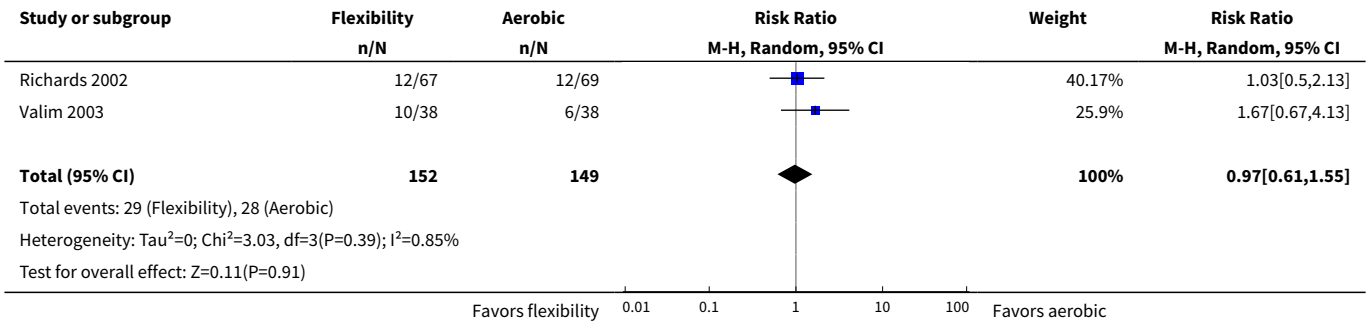
Study or subgroup	Flexibility N	Flexibility Mean(SD)	Aerobic N	Aerobic Mean(SD)	Std. Mean Difference Random, 95% CI	Weight	Std. Mean Difference Random, 95% CI
Matsutani 2012	12	13.9 (5.2)	7	15.5 (3.8)		8.36%	-0.32[-1.26,0.62]
McCain 1988	20	-158.8 (91.2)	18	-165.4 (98.6)		17.24%	0.07[-0.57,0.71]
Richards 2002	67	11 (5)	69	10.2 (5.6)		49.56%	0.15[-0.19,0.49]
Valim 2003	28	14 (4)	32	11 (6)		24.84%	0.57[0.05,1.09]
Total ***	**127**		**126**			**100%**	**0.2[-0.08,0.48]**
Heterogeneity: Tau^2=0.01; Chi^2=3.43, df=3(P=0.33); I^2=12.45%							
Test for overall effect: Z=1.42(P=0.16)							



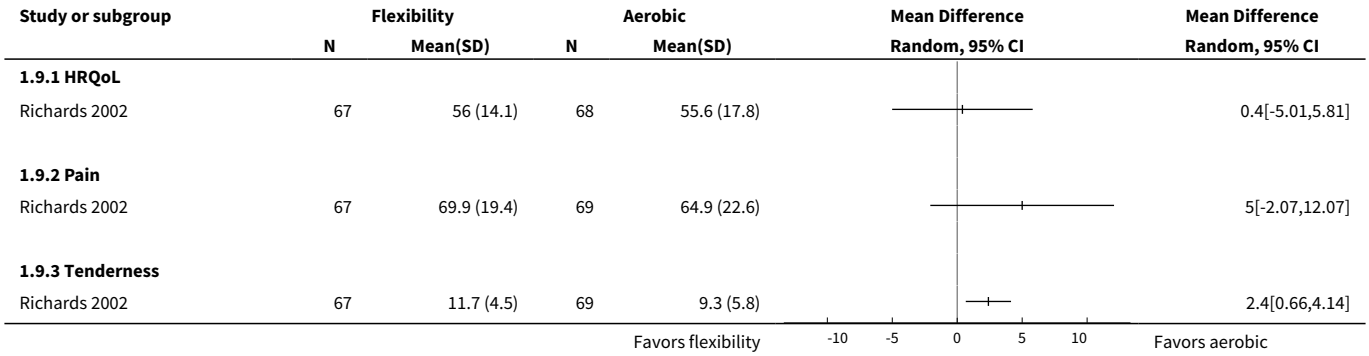
Analysis 1.8. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 8 Withdrawals.

Study or subgroup	Flexibility n/N	Aerobic n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Bressan 2008	0/8	0/7			Not estimable
Matsutani 2012	5/17	8/15		27.76%	0.55[0.23,1.32]
McCain 1988	2/22	2/20		6.17%	0.91[0.14,5.86]





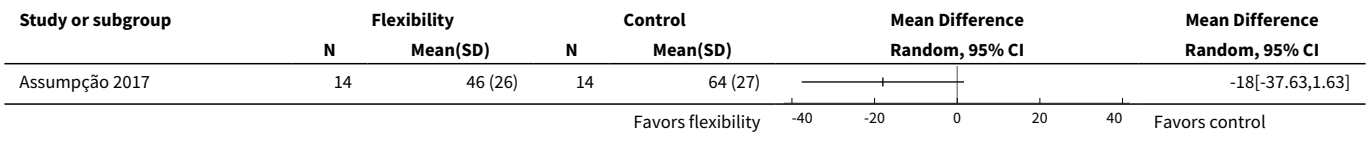
Analysis 1.9. Comparison 1 Flexibility versus aerobic (end of intervention), Outcome 9 Long-term effects.



Comparison 2. Flexibility versus control (end of intervention)

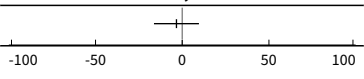
Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 Pain, Intensity, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 Physical function, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
3 Withdrawals	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

Analysis 2.1. Comparison 2 Flexibility versus control (end of intervention), Outcome 1 Pain, Intensity, 0-100, lower is best (end of intervention).



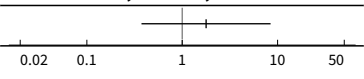
Analysis 2.2. Comparison 2 Flexibility versus control (end of intervention), Outcome 2 Physical function, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Control N	Control Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Assumpção 2017	14	31.7 (17.3)	14	35 (17.7)		-3.33[-16.29,9.63]



Analysis 2.3. Comparison 2 Flexibility versus control (end of intervention), Outcome 3 Withdrawals.

Study or subgroup	Flexibility n/N	Control n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Assumpção 2017	4/18	2/16		1.78[0.37,8.44]

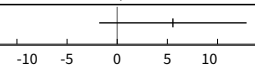


Comparison 3. Flexibility versus resistance (end of intervention)

Outcome or subgroup title	No. of studies	No. of partici-pants	Statistical method	Effect size
1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2 Pain, Intensity, 0-100, lower is best (end of intervention)	3	152	Mean Difference (IV, Random, 95% CI)	1.84 [-4.15, 7.83]
3 Fatigue, 0-100, lower is best (end of intervention)	2	122	Mean Difference (IV, Random, 95% CI)	9.83 [-5.30, 24.97]
4 Physical function, 0-100, lower is best (end of intervention)	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
5 Depression, 0-63, lower is best (end of intervention)	2	122	Mean Difference (IV, Random, 95% CI)	0.47 [-3.40, 4.35]
6 Tenderness, 0-18, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
7 > 30% improvement of pain (end of intervention)	1	30	Odds Ratio (M-H, Random, 95% CI)	0.93 [0.21, 4.11]
8 Withdrawals	3	159	Risk Ratio (M-H, Random, 95% CI)	1.43 [0.77, 2.67]

Analysis 3.1. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Jones 2002	28	43.4 (19.6)	28	37.8 (3.2)		5.55[-1.8,12.9]



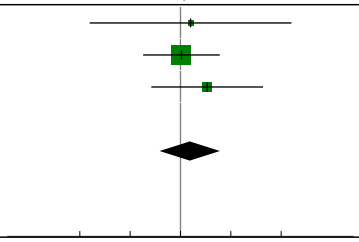
Favors flexibility | Favors resistance

Analysis 3.2. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Assumpção 2017	14	46 (26)	16	44 (30)		8.93%	2[-18.04,22.04]
Gavi 2014	31	57.5 (16.5)	35	57.3 (14.8)		61.9%	0.19[-7.42,7.8]
Jones 2002	28	51.4 (21.7)	28	46.1 (20.6)		29.17%	5.3[-5.79,16.39]
Total ***	**73**		**79**			**100%**	**1.84[-4.15,7.83]**

Heterogeneity: Tau^2=0; Chi^2=0.55, df=2(P=0.76); I^2=0%

Test for overall effect: Z=0.6(P=0.55)



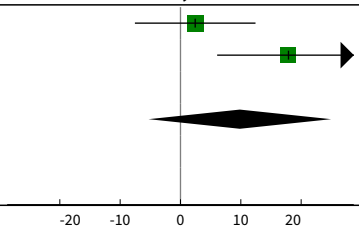
Favors flexibility | Favors resistance

Analysis 3.3. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Gavi 2014	31	61.3 (21.1)	35	58.9 (20.2)		52.17%	2.44[-7.55,12.43]
Jones 2002	28	70 (24.3)	28	52.1 (20.6)		47.83%	17.9[6.08,29.72]
Total ***	**59**		**63**			**100%**	**9.83[-5.3,24.97]**

Heterogeneity: Tau^2=88.32; Chi^2=3.83, df=1(P=0.05); I^2=73.9%

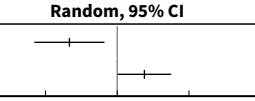
Test for overall effect: Z=1.27(P=0.2)



Favors flexibility | Favors resistance

Analysis 3.4. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 4 Physical function, 0-100, lower is best (end of intervention).

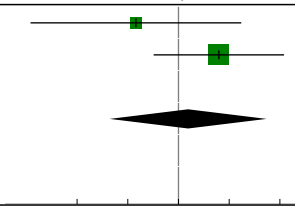
Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Assumpção 2017	14	31.7 (17.3)	16	48.3 (16.7)		-16.66[-28.87,-4.45]
Gavi 2014	31	61.6 (18.9)	35	52.1 (19.8)		9.47[0.13,18.81]



Favors flexibility | Favors resistance

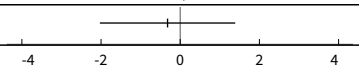
Analysis 3.5. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 5 Depression, 0-63, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Weight	Mean Difference Random, 95% CI
Gavi 2014	31	16.4 (9.5)	35	18.5 (12)		37.04%	-2.1[-7.3,3.1]
Jones 2002	28	10.3 (6.1)	28	8.3 (6.1)		62.96%	1.99[-1.23,5.21]
Total ***	**59**		**63**			**100%**	**0.47[-3.4,4.35]**
Heterogeneity: Tau²=3.51; Chi²=1.72, df=1(P=0.19); I²=41.91%							
Test for overall effect: Z=0.24(P=0.81)							



Analysis 3.6. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 6 Tenderness, 0-18, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Resistance N	Resistance Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
Jones 2002	28	14.7 (3.5)	28	15 (3)		-0.32[-2.03,1.39]



Analysis 3.7. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 7 > 30% improvement of pain (end of intervention).

Study or subgroup	Flexibility n/N	Resistance n/N	Odds Ratio M-H, Random, 95% CI	Weight	Odds Ratio M-H, Random, 95% CI
Assumpção 2017	5/14	6/16		100%	0.93[0.21,4.11]
Total (95% CI)	**14**	**16**		**100%**	**0.93[0.21,4.11]**
Total events: 5 (Flexibility), 6 (Resistance)					
Heterogeneity: Not applicable					
Test for overall effect: Z=0.1(P=0.92)					

Analysis 3.8. Comparison 3 Flexibility versus resistance (end of intervention), Outcome 8 Withdrawals.

Study or subgroup	Flexibility n/N	Resistance n/N	Risk Ratio M-H, Random, 95% CI	Weight	Risk Ratio M-H, Random, 95% CI
Assumpção 2017	4/18	3/19		21.22%	1.41[0.36,5.43]
Gavi 2014	9/31	5/35		40.28%	2.03[0.76,5.42]
Jones 2002	6/28	6/28		38.5%	1[0.37,2.73]
Total (95% CI)	**77**	**82**		**100%**	**1.43[0.77,2.67]**
Total events: 19 (Flexibility), 14 (Resistance)					
Heterogeneity: Tau²=0; Chi²=0.98, df=2(P=0.61); I²=0%					
Test for overall effect: Z=1.13(P=0.26)					

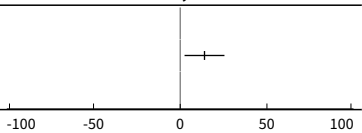
Comparison 4. Flexibility versus other comparators (end of intervention)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention)	3		Mean Difference (IV, Random, 95% CI)	Totals not selected
1.1 Flexibility vs Pilates	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
1.2 Flexibility vs Tai Chi	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
1.3 Flexbility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2 Pain, Intensity, 0-100, lower is best (end of intervention)	3		Mean Difference (IV, Random, 95% CI)	Totals not selected
2.1 Flexibility vs Pilates	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2.2 Flexibility vs Tai Chi	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2.3 Flexibility vs friction massage	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
2.4 Flexibility vs medication	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
3 Fatigue, 0-100, lower is best (end of intervention)	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
3.1 Flexibility vs Tai Chi	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
3.2 Flexibility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
4 Stiffness, 0-100, lower is best (end of intervention)	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
4.1 Flexibility vs Tai Chi	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
4.2 Flexibility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
5 Physical function, 0-100, lower is best (end of intervention)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
5.1 Flexibility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

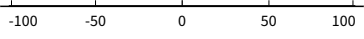
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6 Depression, 0-63, lower is best (end of intervention)	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
6.1 Flexibility vs Tai Chi	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
6.2 Flexibility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
7 Tenderness, 0-18, lower is best (end of intervention)	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
7.1 Flexibility vs Pilates	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
7.2 Flexibility vs aquatics	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
8 Withdrawals	4		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
9 Long-term effects: flexibility vs other comparators	2		Mean Difference (IV, Random, 95% CI)	Totals not selected
9.1 HRQoL	2		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.2 Pain	2		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.3 Fatigue	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.4 Stiffness	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.5 Depression	1		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]
9.6 Tenderness	2		Mean Difference (IV, Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 4.1. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 1 HRQoL, FIQ Total, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.1.1 Flexibility vs Pilates						
Altan 2009	24	77.5 (21.4)	25	63.5 (19.6)		14[2.5,25.5]

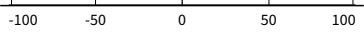


Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.1.2 Flexibility vs Tai Chi						
Calandre 2009	39	55.5 (14.1)	42	51.7 (16.6)		3.8[-2.89,10.49]
4.1.3 Flexbility vs aquatics						
López-Rodríguez 2012	20	69.2 (12.9)	19	52.2 (16.2)		17.07[7.86,26.28]



Analysis 4.2. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 2 Pain, Intensity, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.2.1 Flexibility vs Pilates						
Altan 2009	24	60 (21)	25	41 (17)		19[8.28,29.72]
4.2.2 Flexibility vs Tai Chi						
Calandre 2009	39	66 (21)	42	66 (23)		0[-9.58,9.58]
4.2.3 Flexibility vs friction massage						
Amanollahi 2013	45	24 (29)	45	52 (33)		-28[-40.84,-15.16]
4.2.4 Flexibility vs medication						
Amanollahi 2013	45	24 (29)	39	32 (28)		-8[-20.21,4.21]



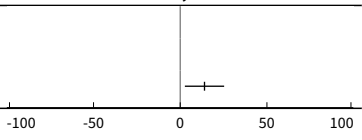
Analysis 4.3. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 3 Fatigue, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.3.1 Flexibility vs Tai Chi						
Calandre 2009	39	79 (20)	42	76 (25)		3[-6.83,12.83]
4.3.2 Flexibility vs aquatics						
López-Rodríguez 2012	20	74.5 (13.9)	19	63.1 (18.5)		11.4[1.09,21.71]

Analysis 4.4. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 4 Stiffness, 0-100, lower is best (end of intervention).

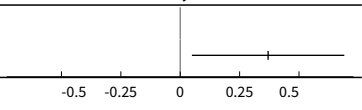
Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.4.1 Flexibility vs Tai Chi						
Calandre 2009	39	68 (24)	42	62 (28)		6[-5.33,17.33]

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.4.2 Flexibility vs aquatics						
López-Rodríguez 2012	20	75.5 (15.7)	19	61.5 (20)		14[2.68,25.32]



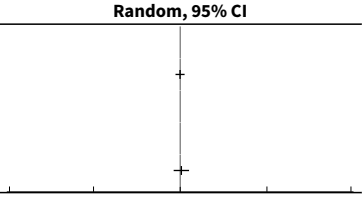
Analysis 4.5. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 5 Physical function, 0-100, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.5.1 Flexibility vs aquatics						
López-Rodríguez 2012	20	1 (0.4)	19	0.6 (0.6)		0.37[0.05,0.69]



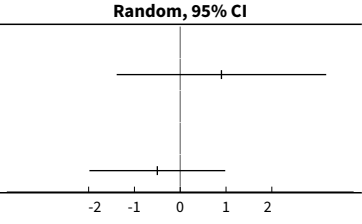
Analysis 4.6. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 6 Depression, 0-63, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.6.1 Flexibility vs Tai Chi						
Calandre 2009	39	9 (5.3)	42	9.1 (6.7)		-0.1[-2.72,2.52]
4.6.2 Flexibility vs aquatics						
López-Rodríguez 2012	20	16.7 (6.7)	19	16.1 (7.4)		0.65[-3.79,5.09]



Analysis 4.7. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 7 Tenderness, 0-18, lower is best (end of intervention).

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.7.1 Flexibility vs Pilates						
Altan 2009	24	14.1 (4.5)	25	13.2 (3.6)		0.9[-1.39,3.19]
4.7.2 Flexibility vs aquatics						
Calandre 2009	39	14.6 (3.1)	42	15.1 (3.7)		-0.5[-1.98,0.98]



Analysis 4.8. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 8 Withdrawals.

Study or subgroup	Flexibility n/N	Other n/N	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
Altan 2009	1/24	0/25		3.12[0.13,73.04]
Amanollahi 2013	6/45	0/45		13[0.75,224.13]
Amanollahi 2013	0/45	0/45		Not estimable
Calandre 2009	5/39	10/42		0.54[0.2,1.44]
López-Rodríguez 2012	15/35	16/35		0.94[0.55,1.59]
		Favors flexibility	0.002 0.1 1 10 500	Favors other

Analysis 4.9. Comparison 4 Flexibility versus other comparators (end of intervention), Outcome 9 Long-term effects: flexibility vs other comparators.

Study or subgroup	Flexibility N	Flexibility Mean(SD)	Other N	Other Mean(SD)	Mean Difference Random, 95% CI	Mean Difference Random, 95% CI
4.9.1 HRQoL						
Altan 2009	24	77.6 (22.2)	25	69.3 (24.7)		8.3[-4.84,21.44]
Calandre 2009	39	57 (13.2)	42	54.7 (14.3)		2.3[-3.69,8.29]
4.9.2 Pain						
Altan 2009	24	65 (21)	25	52 (25)		13[0.09,25.91]
Calandre 2009	39	69 (22)	42	71 (22)		-2[-11.59,7.59]
4.9.3 Fatigue						
Calandre 2009	39	80 (17)	42	78 (18)		2[-5.62,9.62]
4.9.4 Stiffness						
Calandre 2009	39	71 (21)	42	71 (22)		0[-9.37,9.37]
4.9.5 Depression						
Calandre 2009	39	14.3 (8.3)	42	14.6 (10.5)		-0.31[-4.4,3.78]
4.9.6 Tenderness						
Altan 2009	24	14.6 (3.6)	25	13.5 (3.8)		1.1[-0.97,3.17]
Calandre 2009	39	15.1 (3.3)	42	15.1 (3.8)		0[-1.55,1.55]
				Favors flexibility	-20 -10 0 10 20	Favors other

ADDITIONAL TABLES

Table 1. FITT-VP parameters

Author, year, intervention	Frequency, times per week	Length in weeks	Intensity	Time/duration	Session, minutes	Type/mode	Pattern
Flexibility versus control							
Assumpção 2017	2 times/ week	12 weeks	Stretch intensity was increased gradually to the point of moderate discomfort.	30 s	40 min	Supervised program focusing on large muscles (triceps surae, gluteus, ischiotibial, paravertebral, latissimus dorsi, hip adductor, pectoralis)	Not mentioned
Flexibility versus aerobic							
Bressan 2008	1 time/week	8 weeks	Not mentioned	30 s	40 to 45 min	Static muscular stretching of the triceps surae, ischiotibial, gluteal, paravertebral, latissimocondyloideus, pectoral, trapezius, and respiratory muscles. Stretching was performed in dorsal decubitus or sitting.	Performed in a series of 5 repetitions
Matsutani 2012	1 time/week	8 weeks	Not mentioned	30 s	45 min	All exercises emphasized breathing and postural alignment corrections.	For each exercise there were 4 replications, holding the stretch for 30 s on each repetition, followed by 30 s of rest.
McCain 1988	3 times/ week	20 weeks	Not mentioned	Not mentioned	60 min	Exercise consisted of flexibility maneuvers such that sustained heart rate responses greater than 115 beats per minute.	Not mentioned
Richards 2002	2 times/ week	12 weeks	Not mentioned	Not mentioned	60 min	Relaxation and flexibility comprised upper and lower limb stretches and relaxation techniques based on the published regimen by Ost 1987.	Not mentioned
Valim 2003	3 times/ week	20 weeks	Not mentioned	30 s	45 min	Stretching program included 17 exercises using both muscles and joints in a general way, including face, cervical, trunk, and extremities.	Not mentioned
Flexibility versus resistance							

Table 1. FITT-VP parameters *(Continued)*

Assumpção 2017	2 times/ week	12 weeks	Stretch intensity was increased gradually to the point of moderate discomfort.	30 s	40 min	Supervised program focusing on large muscles (triceps surae, gluteus, ischiotibial, paravertebral, latissimus dorsi, hip adductor, pectoralis).	Not mentioned
Gavi 2014	2 times/ week	16 weeks	Not mentioned	30 s	45 min	Stretching program included major muscle groups. Authors reference the stretching protocol used by Valim 2003.	Not mentioned
Jones 2002	2 times/ week	12 weeks	Not mentioned	60 s	60 min	Stretching program included stretches performed in standing, sitting, or lying positions.	Not mentioned
Flexibility versus other							
Altan 2009	3 times/ week	12 weeks	Not mentioned	6 s	60 min	Non-weight bearing stretching of cervical, shoulder, thoracic, lumbar, gluteal leg and crusis muscle	Not mentioned
Amanollahi 2013	3 times/ week	4 weeks	Not mentioned	30 s	Not mentioned	Non-weight bearing stretching of shoulders blade musculature, paraspinal muscles, neck and low back muscle, hamstrings and calf muscles	Each time included 3 repetitions of each stretching exercise
Calandre 2009	3 times/ week	6 weeks	Not mentioned	Not mentioned	60 min	Stretching exercises were performed on muscles over the main body area: cervical, upper and lower groups extremities, and trunk.	Not mentioned
López-Rodríguez 2012	2 times/ week	12 weeks	Not mentioned	Not mentioned	60 min	Flexibility stretching exercises that included global stretches and specific to different muscular areas of the body	Not mentioned

Table 2. Outcome measures used for analysis in the included studies

Outcome	Name of instrument or index/subscale
Health-related quality of life	FIQ Total[1] (0 to 100)
Pain intensity	Current pain (VAS), FIQ pain[1] (VAS), SF-36 bodily pain
Fatigue	FIQ fatigue[1] (0 to 100), SF-36 Vitality (0 to 100)
Stiffness	FIQ stiffness[1] (0 to 100)
Physical function	FIQ physical function[1] (0 to 100), SF-36
Depression	Beck Depression Inventory (0 to 63), FIQ depression[1] (0 to 100)
Tenderness	Tender point count (0 to 18), total myalgic score
Adverse events	Not a standardized instrument or index/narrative information

FIQ: Fibromyalgia Impact Questionnaire; SF-36: 36-item Short Form Health Survey; VAS: visual analogue scale
[1]The revised FIQ scale, Bennett 2009, and any language-translated version of the FIQ (Portuguese version; Assumpção 2017) were considered to be equivalent to the original version of the FIQ (Burckhardt 1991).

Table 3. Detailed description of exercise protocol

Study	Group (naming of the intervention as described by author)	Flexibility	Aerobic	Strength	Other
Altan 2009	Length: 24 weeks 1. HOME EXERCISE 1 h, 3/week RELAXATION/STRETCHING 1 h, 3/week 2. PILATES 1 h, 3/week	1. Muscle groups/exercises: stretching of cervical, shoulder, thoracic, lumbar, gluteal, leg and cruris muscle groups. Holding each stretch for 6 s and relaxed for 4 s 2. None	1. None 2. None	1. None 2. The protocol comprised 9 modules covering postural education, search for neutral position, sitting exercise, antalgic exercises, and breathing education. Equipment: resistance bands and 26-centimeter Pilates balls were used as supportive equipment. The following components were included in the exercises: resistance and stabilization, flexibility and range of motion, proper body alignment, balance, co-ordination, and body awareness. 1-hour program (5 min	1. None 2. None

Table 3. Detailed description of exercise protocol *(Continued)*

				breathing, 10 min warm-up, 35 min conditioning, 10 min cool-down)	
Amanollahi 2013	Length: 4 weeks 1. FLEXIBILITY 3/week 2. MEDICATION 3/day and 1/day 3. FRICTION MASSAGE 3/week	1. Static and non-weight-bearing stretching of shoulders blade musculature, paraspinal muscles, neck and low back muscle, hamstrings and calf muscles. 3 reps with 30 s holds 2. None 3. None	1. None 2. None 3. None	1. None 2. None 3. None	1. None 2. 400 mg ibuprofen (Aria Pharmaceutical Co., Iran) 3 x/day and 25 mg nortriptyline (Darou Pakhsh Pharmaceutical Mfg. Co, Iran) 1/day 3. 3 30-second friction massages using the second and third fingers with a pressure of approximately 0.5 to 1 kg/point on the painful spot so that a mild pallor occurred on the practitioner's nails
Assumpção 2017	Length: 12 weeks 1. FLEXIBILITY 2/week 2. RESISTANCE 2/week	1. Supervised program focusing on large muscles (triceps surae, gluteus, ischiotibial, paravertebral, latissimus dorsi, hip adductor, pectoralis). In early stages 3 reps, from fifth week 4 reps, from ninth week 5 reps; intensity of stretch was gradually increased to point of moderate discomfort and held for 30 s holds for 40 min. 2. None 3. None	1. None 2. None 3. None	1. None 2. Dumbells for upper limbs and shin pads for lower limbs; exercises targeted triceps surae, quadriceps, hip adductors and abductors, hip flexors, elbow flexors and extensors, pectoralis major, and rhomboids.	1. None 2. None 3. None

Table 3. Detailed description of exercise protocol *(Continued)*

	3. CONTROL	2. None 3. None		Duration of 40 min (5 min breathing, 10 min warm-up, 35 min conditioning, 10 min cool-down); first 2 sessions there was no load; 0.5 kg was added each week if participant identified the effort as slightly intense on the Borg Scale (score = 13); 8 reps 3. None	
Bressan 2008	Length: 8 weeks 1. STRETCHING 1/week 2. PHYSICAL CONDITIONING EXERCISES 1/week	1. Static stretches of triceps surae, ischiotibial, gluteal, paravertebral, latissimo-condyloideus, pectoral, trapezius, and respiratory muscles. In addition, stretching at home was recommended. Exercises were performed in a series of 5 repetitions, with 30 s holds for 40 to 45 min. 2. None	1. None 2. Walking for a period of 30 min using a motorized treadmill (5 min warm-up, 25 min walking, 5 min rest). The walking speed was determined at 60% to 75% of the maximum HR, deducting participant's age from 220.	1. None 2. None	1. None 2. None
Calandre 2009	Length: 6 weeks 1. STRETCHING (in water) 1 h, 3/week 2. TAI CHI (in water) 1 h, 3/week	1. Training was done in a pool with water heated at 36 °C and was preceded by a shower with warm water (34.5 °C to 35.5 °C). In order to facilitate the stretching, participants were given 1-meter-long wooden sticks. Stretching was performed over the muscles of main body areas: cervical area, upper and lower extremities, and trunk. 2. None	1. None 2. None	1. None 2. None	1. None 2. Participants were taught the 16 movements which constitute the Tai Chi therapy without the assistance of additional material. Tai Chi is performed standing in shoulder-depth water using a combination of deep breathing and slow, broad movements of the arms,

Table 3. Detailed description of exercise protocol *(Continued)*

					legs, and torso.
Gavi 2014	Length: 16 weeks 1. FLEXIBILITY 45 min, 2/week 2. RESISTANCE TRAINING 45 min, 2/week	1. Stretching program included the major muscles groups. Valim 2003 is referenced for stretching program. 2. None	1. None 2. Resistance training group received supervised progressive training in the standing and sitting positions using weight machines. The intensity was moderate, with an overload of 45% of the estimated 1 RM, calculated based on maximal repetitions. 8 major groups were trained (quadriceps, femoris, hamstrings, biceps brachii, triceps brachii, pectoral, calf, deltoid, and latissimus dorsi) in 12 different exercises, with 3 sets of 12 repetitions (leg press, leg extension, hip flexion, pectoral fly, triceps extension, shoulder flexion, leg curl, calf, pulldown, shoulder abduction, biceps flexion, and shoulder extension).	1. None 2. None	1. None 2. None
Jones 2002	Length: 12 weeks 1. FLEXIBILITY 1 h, 2/week 2. STRENGTH 1 h, 2/week	1. The muscles included in the protocol were gastrocnemius, tibialis anterior, quadriceps, hamstrings, gluteus, abdominals, erector spinae, pectorals, latissimus dorsi, rhomboids, deltoids, biceps, triceps. Static stretch, participant controlled intensity of stretches. 10 min warm-up, 40 min stretching, 10 min cool-down of guided imagery and relaxation 2. Warm-up and cool-down	1 and 2 warm-up	1. None 2. The muscles included in the protocol were gastrocnemius, tibialis anterior, quadriceps, hamstrings, gluteus, abdominals, erector spinae, pectorals, latissimus dorsi and rhomboids, deltoids, biceps, triceps. Equipment used: 1- to 3-pound weights and/or surgical tubing. Concentric/eccentric contractions with minimized work during eccentric phase. Intensity and progression directed by participant. Single set throughout, repetitions progressed from 4 or 5 to 12. Participants encouraged to decrease activity during fibromyalgia flares. 1-hour program including 5 min warm-up, 45 min strengthening, 10 min cool-down	1. None 2. None

Table 3. Detailed description of exercise protocol *(Continued)*

López-Rodríguez 2012	Length: 12 weeks 1. (CONTROL) FLEXIBILITY 1 h, 2/week 2. EXPERIMENTAL GROUP biodanza 1 h, 2/week	1. Flexibility stretching exercises that included global stretches and stretches specific to different muscular areas of the body 2. None	1. None 2. Biodanza in the water with water temperature approximately of 29 °C preceded by a shower at 33 °C to 35 °C, biodanza-type movements like walking, slow movements of upper and lower extremities, cool-down stretching. The duration of the intervention was 60 min (10 min warm-up, 4 min biodanza, 10 min cool-down).	1. None 2. None	1. None 2. None
Matsutani 2012	Length: 8 weeks 1. STRETCHING 45 min, 1/week 2. AEROBIC 30 min, daily	1. Static stretching exercises were performed in a segment of the muscle groups: triceps leg, gluteal, iliopsoas, hamstring, paraspinal, latissimus dorsi, diaphragm, adductor pubic associated with lumbar pelvic movements, trapezius, and major and minor pectoralis. All exercises emphasized breathing and postural alignment. Static stretches held 30 s, repeated 4 times with 30 s rest, progressed from lying to sitting to standing upright or in flexion. Breathing and postural alignment were emphasized. A mirror was used as an aid to the perception of movements of the upper limbs and postural alignment. 2. None	1. None 2. A treadmill walk was performed with intensity defined according to HR, between 60% and 70% HR for age (formula used, HR max = 220−age).	1. None 2. None	1. None 2. None
McCain 1988	Length: 20 weeks 1. FLEXIBILITY 1 h, 3/week 2. AEROBIC EXERCISE 1 h, 3/week	1. Exercises consisted of flexibility maneuvers such that sustained HR responses greater than 115 beats per min were not attained. 2. None	1. None 2. After a 10-minute preliminary warm-up exercise, individuals were subjected to sustained HR elevation training through the use of a bicycle ergometer. Heart rate was maintained in excess of 150 beats per minute for gradually increasing time periods.	1. None 2. None	1. None 2. None
Richards 2002	Length: 12 weeks 1. RELAXATION AND FLEXIBILITY 1 h, 2/week 2. AEROBIC EXERCISE 1 h, 2/week	1. Relaxation and flexibility comprised upper and lower limb stretches and relaxation techniques based on the published regimen by Ost 1987. As the classes proceeded, more techniques were introduced progressing through progressive muscle relaxation, release-only	1. None 2. Exercise therapy comprised an individualized aerobic exercise program, mostly walking on treadmills and cycling on exercise bicycles. Each individual was encouraged to steadily increase the amount of exercise as tolerated.	1. None 2. None	1. None 2. None

Table 3. Detailed description of exercise protocol *(Continued)*

		relaxation and visualization, cue-controlled relaxation, and differential relaxation. 2. None			
Valim 2003	Length: 20 weeks 1. STRETCHING EXERCISE GROUP 45 min, 3/ week 2. AEROBIC EXERCISE GROUP 45 min, 3/ week	1. 17 static exercises using both muscles and joints in a general way, including face, cervical , trunk, and extremities. Exercises chosen to provide flexibility without increasing HR. Each maximum position was sustained for 30 s. 2. None	1. None 2. Exercise group underwent a walking program monitored with frequency meters and supervised by a physiotherapist. The walking speed (training load) was determined by the training HR. Training HR defined as the load beat immediately preceding the one in which the anaerobic threshold occurred. Each training session was preceded by a warm-up period in which participants were instructed to walk freely and slowly for 5 to 10 min. After each session the participants were placed in a circle and performed rhythmic movements, to promote cooling off, for 5 min.	1. None 2. None	1. None 2. None

HR: heart rate; RM: maximum repetition; Max: maximum

Table 4. Congruence with 2013 ACSM flexibility criteria for healthy adults *(Continued)*

Author, year	Met ACSM 2013 criteria					
	Frequency	**Intensity**	**Time**	**Type**	**Volume**	**Pattern**
	2 to 3 d/ week with daily being most effective	**Stretch to the point of feeling tightness or slight discomfort**	**10 s to 30 s**	**A series of flexibility exercises for each of the major muscle-tendon units**	**60 s of total stretching time for each flexibility exercise**	**2 to 4 repetitions**
Altan 2009	Yes	Unclear	No	Yes	Unclear	Unclear
Amanollahi 2013	Yes	Unclear	Yes	Yes	Yes	Yes
Assumpção 2017	Yes	Yes	Yes	Yes	Unclear	Unclear
Bressan 2008	No	Unclear	Yes	Yes	Yes	Yes
Calandre 2009	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Gavi 2014	Yes	Unclear	Yes	Yes	Unclear	Unclear
Jones 2002	Yes	Unclear	Yes	Yes	Unclear	Unclear

Table 4. Congruence with 2013 ACSM flexibility criteria for healthy adults *(Continued)*

López-Rodríguez 2012	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Matsutani 2012	No	Unclear	Yes	Yes	Yes	Yes
McCain 1988	Yes	Unclear	Unclear	Unclear	Unclear	Unclear
Richards 2002	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Valim 2003	Yes	Unclear	Yes	Yes	Unclear	Unclear

Table 5. Quality of evidence—GRADE assessment: long-term effects of flexibility exercise training versus aerobic exercise training

Certainty assessment						№ of participants		Certainty	Importance
№ of studies and study design	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Flexibility**	**Aerobic (end of intervention)**		
HRQoL (follow-up 36 weeks after end of intervention; assessed with FIQ Total 0 to 100, lower is best)									
1 randomized trial	Serious[a]	Not serious	Very serious[b]	Serious[c]	None	67	68	⊕◯◯◯ VERY LOW	CRITICAL
Pain intensity (follow-up 36 weeks after end of intervention; assessed with VAS 0 to 100, lower is best)									
1 randomized trial	Serious[a]	Not serious	Very serious[b]	Serious[c]	None	67	69	⊕◯◯◯ VERY LOW	CRITICAL
Fatigue, stiffness, and physical function: not measured									
Withdrawals, adverse events: not reported									

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life; VAS: visual analogue scale
[a]Downgraded one level for selection bias.
[b]Downgraded two levels because flexibility was used as a proxy (i.e. flexibility exercise was used along with relaxation as the control in the study).
[c]Downgraded one level for imprecision (sample size lower than 400 rule-of-thumb).

Table 6. Quality of evidence—GRADE assessment: flexibility intervention versus control *(Continued)*

№ of studies and study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	№ of participants		Certainty	Importance
						Flexibility	**Control (end of intervention)**		
Pain, intensity, 0 to 100, lower is best (end of intervention)									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[b]	None	14	14	⊕⊕◯◯ LOW	CRITICAL
Physical function, 0 to 100, lower is best (end of intervention)									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[b]	None	14	14	⊕⊕◯◯ LOW	CRITICAL

Table 6. Quality of evidence—GRADE assessment: flexibility intervention versus control *(Continued)*

Withdrawals									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[b]	None	4/18 (22.2%)	2/16 (12.5%)	⨁⨁◯◯ LOW	IMPORTANT
HRQoL, fatigue, and stiffness: data were described as skewed, thus were not used									
Adverse events: not measured/reported for either group									

HRQoL: health-related quality of life
[a]Downgraded one level because of selection and performance bias.
[b]Downgraded one level because of imprecision (sample size lower than 400 rule-of-thumb).

Table 7. Quality of evidence—GRADE assessment: flexibility intervention versus resistance training intervention

Certainty assessment						№ of participants		Certainty	Importance
№ of studies and study design	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Flexibility**	**Resistance (at end of intervention)**		
HRQoL, FIQ Total, 0 to 100, lower is best (end of intervention)									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[b]	None	28	28	⨁⨁◯◯ LOW	CRITICAL
Pain, intensity, 0 to 100, lower is best (end of intervention)									
3 randomized trials	Serious[a]	Not serious	Not serious	Serious[b]	None	73	79	⨁⨁◯◯ LOW	CRITICAL
Fatigue, 0 to 100, lower is best (end of intervention)									
2 randomized trials	Very serious[c]	Serious[d]	Not serious	Serious[b]	None	59	63	⨁◯◯◯ VERY LOW	IMPORTANT
Physical function, 0 to 100, lower is best (end of intervention)									
2 randomized trials	Serious[a]	Very serious[e]	Not serious	Serious[b]	None	45	51	⨁◯◯◯ VERY LOW	IMPORTANT

Table 7. Quality of evidence—GRADE assessment: flexibility intervention versus resistance training intervention *(Continued)*

> 30% improvement of pain (end of intervention)									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[b]	None	5/14 (35.7%)	6/16 (37.5%)	⊕⊕◯◯ LOW	IMPORTANT
Withdrawals									
3 randomized trials	Serious[a]	Not serious	Not serious	Serious[b]	None	19/77 (24.7%)	14/82 (17.1%)	⊕⊕◯◯ LOW	IMPORTANT
Stiffness: not measured									
Adverse events: not measured/reported for flexibility training group For resistance training group, "one subject in the resistance group interrupted participation in the study because of worsening pain" (page 13 of 22)									

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life
[a]Downgraded one level because of selection and performance bias.
[b]Downgraded one level for imprecision (sample size lower than 400 rule-of-thumb).
[c]Downgraded two levels because of selection and performance bias.
[d]Downgraded one level for inconsistency.
[e]Considerable heterogeneity (I^2 = 91%).

Table 8. Quality of evidence—GRADE assessment: flexibility intervention versus other comparators

Certainty assessment						№ of participants		Certainty	Importance
№ of studies and study design	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Flexibility**	**Other comparators (end of intervention)**		
HRQoL, FIQ Total, 0 to 100, lower is best (end of intervention)									
3 randomized trials	Serious[a]	Serious[b]	Not serious	Serious[c]	Studies not pooled	83	86	⊕◯◯◯ VERY LOW	CRITICAL
Pain, intensity, 0 to 100, lower is best (end of intervention)									
3 randomized trials	Serious[a]	Serious[b]	Not serious	Serious[c]	Studies not pooled	153	151	⊕◯◯◯ VERY LOW	CRITICAL

Table 8. Quality of evidence—GRADE assessment: flexibility intervention versus other comparators *(Continued)*

Fatigue, 0 to 100, lower is best (end of intervention)									
2 randomized trials	Serious[a]	Serious[b]	Not serious	Serious[c]	Studies not pooled	59	61	⊕◯◯◯ VERY LOW	IMPOR-TANT
Stiffness, 0 to 100, lower is best (end of intervention)									
2 randomized trials	Serious[a]	Serious[b]	Not serious	Serious[c]	Studies not pooled	59	61	⊕◯◯◯ VERY LOW	IMPOR-TANT
Physical function, 0 to 100, lower is best (end of intervention)									
1 randomized trial	Serious[a]	Not serious	Not serious	Serious[c]	1 study	20	19	⊕⊕◯◯ LOW	IMPOR-TANT
Withdrawals									
4 randomized trials	Serious[a]	Serious	Not serious	Serious[c]		27/188 (14.4%)	26/192 (13.5%)	⊕◯◯◯ VERY LOW	IMPOR-TANT
Adverse events: not reported for flexibility group In the medication arm, 5 participants who received ibuprofen and 1 participant who received nortriptyline experienced side effect (from translated version of article).									

FIQ: Fibromyalgia Impact Questionnaire; HRQoL: health-related quality of life
[a]Downgraded one level because of selection and performance bias.
[b]Interventions not consistent across studies.
[c]Downgraded one level for imprecision (sample size lower than 400 rule-of-thumb).

APPENDICES

Appendix 1. Glossary of terms

Term	Meaning
Aerobic exercise training	Aerobic exercise training primarily affects the circulatory system and the respiratory system. Following aerobic exercise training, the heart pumps out more blood per beat and there are more capillaries available to transfer this blood to the working muscles and to the lungs. In addition, the lungs become more efficient at moving air in and out and in transferring oxygen into the blood and removing carbon dioxide. As a result of these improvements in heart and lung function, people have an increased total work capacity, and they can do a higher rate of work at a given submaximal level (ACSM 2013).
Ballistic stretching	Exercise characterized by repeated bouncing
Cardiorespiratory fitness	The ability of the cardiovascular and respiratory systems to supply oxygen to muscles during sustained physical activity
Companion study	In keeping with Rosenthal's recommendations (Rosenthal 1995), publications referring to the same primary study were called companions.
Complex activities	An intervention comprising multiple components which interact to produce change. Complexity may also relate to the difficulty of behaviors targeted by interventions, the number of organizational levels targeted, or the range of outcomes.
Concomitant	Existing or occurring with something else
Confidence interval	A range of values so defined that there is a specified probability that the value of a parameter lies within it
Dolorimetry	A method of measuring intensity of pain perception in degrees ranging from unpleasant to unbearable by using heat applied to the skin as a gauge
Efficacy	The ability to produce a desired or intended result
Exercise	Physical activity that is planned, structured, and repetitive and that has as a final or intermediate objective the improvement or maintenance of physical fitness (Caspersen 1985)
Exercise training	Program that is designed to meet individual health and physical fitness goals; a single exercise session should include warm-up, stretching, conditioning, and cool-down components. The rate of progression depends on the individual's health status and exercise tolerance.
Exacerbation	Worsening of signs and symptoms
Flexibility exercise training	Flexibility exercise training targets the major muscle tendon units of the shoulder girdle, chest, neck, trunk, lower back, hips, posterior and anterior legs, and ankles. There are several types of flexibility exercises that can improve range of motion such as ballistic methods, or 'bouncing' stretches, dynamic or slow movement stretching, static stretching, active or passive stretching, and proprioceptive neuromuscular facilitation (PNF) (ACSM 2013).
Maximum heart rate (HRmax)	The highest number of beats per minute the heart can reach during maximum physical exertion. It is unique to each individual and depends on hereditary factors and age.
Muscle strength	The amount of force a muscle can generate

(Continued)

Myalgic score	A measure used to monitor the patient's condition through patient's report of pain or tenderness
Non-pharmacological or non-pharmacologic	Treatment that does not include medication
Neuromuskuloskeletal	Including components of the nervous system (e.g. peripheral nerves and the brain), the muscular system (muscles and tendons), and the skeletal system (bones)
PEDro scale	A measure of the methodological quality of clinical trials. The scale considers two aspects of trial quality, namely the 'believability' (or 'internal validity') of the trial, and whether the trial contains sufficient statistical information to make it interpretable.
Proprioceptive neuromuscular facilitation (PNF)	A rehabilitation technique used to stimulate the neuromuscular system in an effort to excite proprioceptors (sensory organs in muscles, tendons, bones, and joints) in order to produce a desired movement
Physical activity	Any bodily movement produced by skeletal muscles that results in energy expenditure above resting (basal) levels. Physical activity broadly encompasses exercise, sports, and physical activities done as part of daily living, occupation, leisure, and active transportation (Caspersen 1985).
Physical fitness	The ability to carry out daily tasks with vigour and alertness, without undue fatigue and with ample energy to enjoy leisure pursuits and to meet unforeseen emergencies. Physical fitness is operationalised as a set of measurable health and skill-related attributes.
Physical function	The capacity of an individual to carry out the physical activities of daily living. Physical function reflects motor function and control, physical fitness, and habitual physical activity and is an independent predictor of functional independence, disability, and morbidity.
Physiology	The branch of biology dealing with the functions and activities of living organisms and their parts, including all physical and chemical processes
Prevalence	Rate of occurrence of a condition, usually expressed on a per-year basis
Pseudorandomization	A pseudorandom process is a process that appears to be random but is not. Pseudorandom sequences typically exhibit statistical randomness while being generated by an entirely deterministic causal process.
Resistance exercise training	Resistance training can take several forms, producing more strength, more power, or more endurance in the muscles. The effects of resistance training are in the muscles and their neuromuscular effectors (ACSM 2013).
Skewness	Not every distribution of data is symmetric—sets of data that are not symmetric are said to be 'asymmetrical.' The measure of how asymmetrical a distribution can be is called 'skewness.'
Sleep disturbance	A score derived from a questionnaire that measures sleep quantity and quality. The Medical Outcomes Survey Sleep Scale measures six dimensions of sleep (initiation, staying asleep, quantity, adequacy, drowsiness, shortness of breath, snoring).
Somatic comorbidities	Conditions of the body related to a disease
Symptoms	A patient's perceptions of an 'abnormal' physical, emotional, or cognitive state
Tenderness	Pain evoked by tactile pressure on the skin surface
Tender points	Anatomic locations used to identify fibromyalgia. The deep diffuse muscular pain is localized to a number of reproducible (from patient to patient) areas that are tender when palpated. Tender points differ from trigger points, in that pain does not radiate to referred areas.

(Continued)

Trial registry	An official platform and catalogue for registering a clinical trial

Appendix 2. 2013 ACSM recommendations for flexibility exercises in healthy adults

FITT-VP	Evidence-based recommendations
Frequency	≥ 2 to 3 days/week, with daily being most effective
Intensity	Stretch to the point of feeling tightness or slight discomfort
Time	Holding a static stretch for 10 to 30 s is recommended for most adults. In older individuals, holding a stretch for 30 to 60 s may confer greater benefit. For proprioceptive neuromuscular facilitation (PNF) stretching, a 3 to 6 s light-to-moderate contraction (e.g. 20% to 75% of maximum voluntary contraction) followed by a 10 to 30 s assisted stretch is desirable.
Type	A series of flexibility exercises for each of the major muscle-tendon units is recommended. Static flexibility (i.e. active or passive), dynamic flexibility, ballistic flexibility, and PNF are each effective.
Volume	A reasonable target is to perform 60 s of total stretching time for each flexibility exercise.
Pattern	Repetition of each flexibility exercise 2 to 4 times is recommended. Flexibilitiy exercise is most effective when the muscle is warmed through light-to-moderate aerobic activity or passively through external methods such as moist heat pack or hot packs.
Progression	Methods for optimal progression are unknown.

Appendix 3. MEDLINE (Ovid) search strategy

1 Fibromyalgia/
2 fibromyalgi$.tw.
3 fibrositis.tw.
4 or/1-3
5 exp Exercise/
6 Physical Exertion/
7 Physical Fitness/
8 exp Physical Endurance/
9 exp Sports/
10 Pliability/
11 exertion$.tw.
12 exercis$.tw.
13 sport$.tw.
14 ((physical or motion) adj5 (fitness or therapy or therapies)).tw.
15 (physical$ adj2 endur$).mp.
16 manipulat$.tw.
17 (skate$ or skating).tw.
18 jog$.tw.
19 swim$.tw.
20 bicycl$.tw.

21 (cycle$ or cycling).tw.
22 walk$.tw.
23 (row or rows or rowing).tw.
24 weight train$.tw.
25 muscle strength$.tw.
26 exp Yoga/
27 yoga.tw.
28 exp Tai Ji/
29 tai chi.tw.
30 ai chi.tw.
31 exp Vibration/
32 vibration.tw.
33 pilates.tw.
34 or/5-33
35 4 and 34

Appendix 4. Embase (Ovid) search strategy

1 FIBROMYALGIA/
2 fibromyalgi$.tw.
3 fibrositis.tw.
4 or/1-3
5 exp exercise/
6 fitness/
7 exercise tolerance/
8 exp sport/
9 pliability/
10 exertion$.tw.
11 exercis$.tw.
12 sport$.tw.
13 ((physical or motion) adj5 (fitness or therapy or therapies)).tw.
14 (physical$ adj2 endur$).tw.
15 manipulat$.tw.
16 (skate$ or skating).tw.
17 jog$.tw.
18 swim$.tw.
19 bicycl$.tw.
20 (cycle$ or cycling).tw.
21 walk$.tw.
22 (row or rows or rowing).tw.
23 weight train$.tw.
24 muscle strength$.tw.
25 or/5-24
26 4 and 25
27 (random$ or placebo$).ti,ab.
28 ((single$ or double$ or triple$ or treble$) and (blind$ or mask$)).ti,ab.
29 controlled clinical trial$.ti,ab.
30 RETRACTED ARTICLE/
31 or/27-30
32 (animal$ not human$).sh,hw.
33 31 not 32
34 26 and 33

Appendix 5. CINAHL (EBSCO) search strategy

S38	S36 AND S37
S37	S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35

(Continued)

S36	S1 OR S2 OR S3
S35	TX vibration
S34	(MH "Vibration")
S33	(MH "Pilates") OR "pilates"
S32	TX pilates
S31	TX tai ji
S30	(MM "Tai Chi")
S29	TX tai chi
S28	TX yoga
S27	(MH "Yoga") OR (MH "Yoga Pose")
S26	TI manipulat* OR AB manipulat*
S25	TI muscle strength* OR AB muscle strength*
S24	TI weight train* OR AB weight train*
S23	TI (row or rows or rowing) OR AB (row or rows or rowing)
S22	TI walk* OR AB walk*
S21	TI (cycl* or cycling) OR AB (cycl* or cycling)
S20	TI bicycl* OR AB bicycl*
S19	TI swim* OR AB swim*
S18	TI jog* OR AB jog*
S17	TI (skate* or skating) OR AB (skate* or skating)
S16	TI physical* N2 endur* OR AB physical* N2 endur*
S15	TI motion N5 fitness OR AB motion N5 fitness OR TI motion N5 therapy OR AB motion N5 therapy OR TI motion N5 therapies OR AB motion N5 therapies
S14	TI physical N5 fitness OR AB physical N5 fitness OR TI physical N5 therapy OR AB physical N5 therapy OR TI physical N5 therapies OR AB physical N5 therapies
S13	TI sport* OR AB sport*
S12	TI exercis* OR AB exercis*
S11	TI exertion* OR AB exertion*
S10	(MH "Physical Endurance+")

(Continued)

S9	(MH "Pliability")
S8	(MH "Sports+")
S7	(MH "Exercise Test+")
S6	(MH "Physical Fitness")
S5	(MH "Exertion+")
S4	(MH "Exercise+")
S3	AB fibrositis OR TI fibrositis
S2	TI fibromyalgia OR AB fibromyalgia
S1	(MH "Fibromyalgia")

Appendix 6. Cochrane Library (Wiley) search strategy

#1MeSH descriptor: [Exercise] explode all trees
#2MeSH descriptor: [Exercise Therapy] explode all trees
#3MeSH descriptor: [Physical Therapy Modalities] explode all trees
#4MeSH descriptor: [Physical Fitness] explode all trees
#5MeSH descriptor: [Exercise Tolerance] explode all trees
#6MeSH descriptor: [Sports] explode all trees
#7MeSH descriptor: [Pliability] explode all trees
#8MeSH descriptor: [Pliability] explode all trees
#9MeSH descriptor: [Motion] explode all trees
#10MeSH descriptor: [Physical Endurance] explode all trees
#11exercise:ti,ab
#12swim:ti,ab
#13skate:ti,ab
#14jog:ti,ab
#15bike:ti,ab
#16cycle:ti,ab
#17walk:ti,ab
#18row:ti,ab
#19weight train:ti,ab
#20muscle strength:ti,ab
#21#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20
#22MeSH descriptor: [Fibromyalgia] explode all trees
#23fibromyalgia:ti,ab
#24 #22 or #23
#25 #21 and #24

Appendix 7. AMED (Ovid) search strategy

1 Fibromyalgia/
2 fibromyalgi$.tw.
3 fibrositis.tw.
4 or/1-3
5 exp exercise/
6 physical fitness/
7 exp physical endurance/
8 exp sports/
9 Pliability/
10 exertion$.tw.

11 exercis$.tw.
12 sport$.tw.
13 ((physical or motion) adj5 (fitness or therapy or therapies)).tw.
14 (physical$ adj2 endur$).tw.
15 manipulat$.tw.
16 (skate$ or skating).tw.
17 jog$.tw.
18 swim$.tw.
19 bicycl$.tw.
20 (cycle$ or cycling).tw.
21 walk$.tw.
22 (row or rows or rowing).tw.
23 weight train$.tw.
24 muscle strength$.tw.
25 exp pilates/
26 exp yoga/
27 Tai chi/
28 tai ji.tw.
29 yoga.tw.
30 (hatha or kundalini or ashtanga or bikram).tw.
31 pilates.tw.
32 exp exercise therapy/
33 or/5-32
34 4 and 33

Appendix 8. ProQuest Thesis and Disseration Abstract Global search strategy

Terms searched fibromyalg* or fibrositis (in citation or abstract)

Appendix 9. PEDro search strategy

PEDro was searched using 6 combinations:

fibromyalg* AND fitness training

fibromyalg* AND strength training

fibromyalg* AND exercise

fibrositis AND fitness training

fibrositis AND strength training

fibrositis AND exercise

Appendix 10. ClinicalTrials.gov search strategy

Terms searched: fibromyalg* or fibrositis

Appendix 11. WHO International Clinical Trials Registry Platform (ICTRP) search strategy

Terms searched: fibromyalg* or fibrositis

Appendix 12. Selection criteria

Level One screen:

Based on title and abstract of the report:

1. Does the study deal exclusively with fibromyalgia? No—exclude, Yes or uncertain—go to step two
2. Does it include exercise? No—exclude, Yes or uncertain—go to step three
3. Does the study deal exclusively with adults? No—exclude, Yes or uncertain—go to step four
4. Is it a randomized controlled trial (RCT)? No—exclude, Yes or uncertain—include

Level Two screen:

Based on the full text or protocol of the report:

1. Does the study deal exclusively with fibromyalgia? No—exclude, Yes—go to step three, Uncertain—add to list of questions for study author and proceed to step three
2. Is the diagnosis of fibromyalgia based on published criteria? No—exclude, Yes—go to step three, Uncertain—add to list of questions for study author and proceed to step three
3. Does the study deal exclusively with adults? No—exclude, Yes—go to step three, Uncertain—add to list of questions for study author and proceed to step three
4. Is it an RCT? (the study uses terms such as 'random,' 'randomized,' 'RCT,' or 'randomization' to describe the study design or assignment of participants to groups) No—exclude, Yes—go to step three, Uncertain—add to list of questions for study author and proceed to step three
5. Does it include exercise (the study involves at least one intervention that includes exercise)? No—exclude, Yes—go to step three, Uncertain—add to list of questions for study author and proceed to step three
6. (for full text) Are between-group data provided for the outcomes? No (the study contains ONLY fibromyalgia, or results are reported such that effects on fibromyalgia cannot be isolated—exclude, Yes—include the study, Yes but uncertain about one or more of steps—reserve judgement until study authors are contacted

Level Three screen (classification of the study using team's intervention listing):

1. Classification of design
2. Number of interventions
3. Type of comparisons:
   - Head-to-head comparison?
   - Exercise to control?
   - Composite to control?

1. Control group
   - Classify type of control

1. Exercise
   - Enter types of exercise interventions used in the study
   - Complete naming of the intervention groups

Appendix 13. Busch 2002 search strategy

Process	Particulars
Databases used	MEDLINE (1966-12/2000), CINAHL (1982-12/2000), HealthSTAR (1990-12/2000), Sports Discus (1975-12/2000), Embase (1974 to 12/2000), the Cochrane Controlled Trials Register (2000, Issue 4)
Adjunctive search methods	Reference lists from identified articles, meta-analyses, and reviews of all types of treatment for fibromyalgia were reviewed independently by two review authors and all promising references were scrutinized. We searched without language restriction and translated all non-English studies that were initially identified as possibly meeting the inclusion criteria
Search strategy used for MEDLINE	Search strategy on SilverPlatter v3.0 for Windows
	1 "Fibromyalgia"/ all subheadings 2 fibromyalgia 3 fibrositis 4 fibromyalgia or fibrositis 5 #1 or #4 6 explode "Exertion"/all subheadings 7 "Physical-Fitness"/all subheadings 8 explode "Physical-Therapy"/all subheadings 9 "Exercise-Test"/all subheadings 10 "Exercise-Tolerance"/all subheadings 11 explode "Sports"/all subheadings 12 "Pliability"/all subheadings 13 #6 or #7 or #8 or #9 or #10 or #11 or #12

(Continued)

14 exertion*
15 exercis*
16 physical 17 motion
18 fitness
19 therapy
20 therapies
21 (physical or motion) near (fitness or therapy or therapies)
22 physical
23 endurance
24 physical near endurance
25 manipulation*
26 skating
27 running
28 jogging
29 swimming
30 bicycling
31 cycling
32 walking
33 rowing
34 weight
35 training
36 muscle
37 strengthening
38 skating or running or jogging or swimming or bicycling or cycling or walking or rowing or weight training or muscle strengthening
39 #13 or #14 or #15 or #21 or #24 or #25 or #38
40 #5 and #39
41 explode "Research-Design"/all subheadings
42 explode "Clinical-Trials"/all subheadings
43 #41 or #42
44 #40 and #43
45 PT = "CLINICAL-TRIAL"
46 #40 and (PT = "CLINICAL-TRIAL")
47 #44 or #46

HISTORY

Review first published: Issue 9, 2019

Date	Event	Description
14 June 2008	Amended	Converted to new review format. CMSG ID C036-R
17 August 2007	New citation required and conclusions have changed	Substantive amendment. See published notes for details.

CONTRIBUTIONS OF AUTHORS

SYK: has been responsible for leading the flexibility exercise training review team, participating in discussion regarding review design and methods, screening studies for eligibility for inclusion in the review, supporting data extraction, performing methodological and statistical analysis, drafting and reviewing the manuscript. Read and approved the final manuscript.

AJB: has been responsible for co-ordinating the fibromyalgia and exercise team, designing and reviewing the review protocol, performing the literature search, assessing eligibility of studies, supporting data extraction, performing methodological and statistical analysis, reviewing the manuscript. Read and approved the final manuscript.

TJO: contributed to data extraction and 'Risk of bias' assessment, supported interpretation of study findings, and contributed to reviewing of drafts. Read and approved the final manuscript.

CLS: has been responsible for screening studies for eligibility, supporting data extraction, providing expert opinion on exercise physiology; contributed to editing of the 'Characteristics of included studies' section, assessing risk of bias, writing and reviewing the manuscript. Read and approved the final manuscript.

IV: supported data extraction, 'Risk of bias' assessment; contributed to editing of the 'Characteristics of included studies' section; wrote and reviewed drafts; and approved the final draft of the manuscript.

CB: performed literature searches; participated in discussion regarding methods; worked on Plain language summary, flow chart, searches to support Background and Discussion sections of the manuscript; assisted with writing and reviewing of the manuscript. Read and approved the final manuscript.

SMG: supported data extraction, interpretation of study findings, 'Risk of bias' assessment; wrote and reviewed drafts of the manuscript; contributed to creation of the 'Summary of findings' table. Read and approved the final manuscript.

HJAF: collaborated on data extraction, provided content expertise on exercise physiology, supported interpretation of study findings, and contributed to writing and reviewing of drafts. Read and approved the final manuscript.

JB: has been responsible for assessing eligibility of studies, supporting data extraction, contributing to methodological and statistical analysis, supporting interpretation of study findings, assessing risk of bias, and reviewing the manuscript. Read and approved the final manuscript.

DECLARATIONS OF INTEREST

We confirm that any present or past affiliations or other involvement in any organization or entity with an interest in the review that might lead me/us to have a real or perceived conflict of interest is listed below.

- Soo Y Kim: none known.
- Angela J Busch: none known.
- Tom J Overend: none known.
- Candice L Schachter: none known.
- Ina van der Spuy: I am employed at the School of Rehabilitation Science, University of Saskatchewan.
- Catherine Boden: none known.
- Suelen M Góes: none known.
- Heather JA Foulds: none known.
- Julia Bidonde: none known.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

None.

NOTES

We prepared no new protocol for this review. This review presents a major update of previous reviews completed in 2002 and 2007 on exercise for fibromyalgia. Although this is an update, review team members ran the search from inception of the databases, screened studies, and extracted data from all included studies. Given growth in the literature, we have split this review into several reviews (ie, resistance, aquatic, mixed, aerobic, flexibility, and whole body vibration). Differences between the 2007 review and this update include the following.

- Further refinement of the definition of what constitutes an flexibility exercise training intervention as required to provide clear criteria on the types of interventions that would and would not meet criteria for inclusion in this review.
- Revisions to search terms, databases, and registries (see Appendix 13).
- Changes in membership of the review team (new review authors and two consumers added).
- Use of the "Risk of bias" tool (Higgins 2011) instead of van Tulder 2003 and Jadad 1996 to assess certainty of the evidence.
- Outcomes presented to facilitate standardization of outcomes between reviews on fibromyalgia within Cochrane.
- Adverse events and withdrawals for groups added to major outcomes to reflect other important potential harmful outcomes of flexibility exercise interventions.
- Revisions to Cochrane methods described in version 5.1.0 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), including addition of the Plain Language Summary, the "Summary of findings" table, and GRADEpro to judge evidence certainty.

- Use of electronic data screening (Covidence) extraction methods (Google Docs) as opposed to paper-based methods used in earlier versions of the review
- Post hoc sensitivity analysis.
- Compliance with MECIR standards.

INDEX TERMS

Medical Subject Headings (MeSH)

*Quality of Life; Exercise; Exercise Therapy [*methods]; Fatigue [*therapy]; Fibromyalgia [physiopathology] [*therapy]; Pain Measurement; Randomized Controlled Trials as Topic; Resistance Training; Treatment Outcome

MeSH check words

Humans