



Cochrane
Library

Cochrane Database of Systematic Reviews

Exercise for improving outcomes after osteoporotic vertebral fracture (Review)

Gibbs JC, MacIntyre NJ, Ponzano M, Templeton JA, Thabane L, Papaioannou A, Giangregorio LM

Gibbs JC, MacIntyre NJ, Ponzano M, Templeton JA, Thabane L, Papaioannou A, Giangregorio LM.
Exercise for improving outcomes after osteoporotic vertebral fracture.
Cochrane Database of Systematic Reviews 2019, Issue 7. Art. No.: CD008618.
DOI: [10.1002/14651858.CD008618.pub3](https://doi.org/10.1002/14651858.CD008618.pub3).

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
SUMMARY OF FINDINGS	4
BACKGROUND	7
OBJECTIVES	7
METHODS	8
RESULTS	10
Figure 1.	11
Figure 2.	15
Figure 3.	16
DISCUSSION	21
AUTHORS' CONCLUSIONS	23
ACKNOWLEDGEMENTS	24
REFERENCES	25
CHARACTERISTICS OF STUDIES	30
DATA AND ANALYSES	52
Analysis 1.1. Comparison 1 Exercise versus Control (After 52 weeks), Outcome 1 Incident Fragility Fracture.	52
Analysis 1.2. Comparison 1 Exercise versus Control (After 52 weeks), Outcome 2 Incident Falls.	52
Analysis 2.1. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 1 Maximal Walking Speed over 20 meters.	53
Analysis 2.2. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 2 Timed Up and Go.	53
Analysis 2.3. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 3 QUALEFFO-41 Physical Function Score.	53
Analysis 2.4. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 4 QUALEFFO-41 Total Score.	54
ADDITIONAL TABLES	54
APPENDICES	56
WHAT'S NEW	63
HISTORY	64
CONTRIBUTIONS OF AUTHORS	64
DECLARATIONS OF INTEREST	64
SOURCES OF SUPPORT	64
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	65
INDEX TERMS	65

[Intervention Review]

Exercise for improving outcomes after osteoporotic vertebral fracture

Jenna C Gibbs¹, Norma J MacIntyre², Matteo Ponzano³, Jeffrey Alan Templeton³, Lehana Thabane⁴, Alexandra Papaioannou⁵, Lora M Giangregorio^{3,6}

¹Department of Kinesiology and Physical Education, McGill University, Montreal, Canada. ²School of Rehabilitation Science, McMaster University, Hamilton, Canada. ³Department of Kinesiology, University of Waterloo, Waterloo, Canada. ⁴Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada. ⁵Department of Medicine, McMaster University, Hamilton, Canada. ⁶Schlegel-UW Research Institute for Aging, Waterloo, Canada

Contact address: Lora M Giangregorio, Department of Kinesiology, University of Waterloo, 200 University Ave W, Waterloo, ON, N2L 3G1, Canada. lora.giangregorio@uwaterloo.ca.

Editorial group: Cochrane Musculoskeletal Group.

Publication status and date: New search for studies and content updated (conclusions changed), published in Issue 7, 2019.

Citation: Gibbs JC, MacIntyre NJ, Ponzano M, Templeton JA, Thabane L, Papaioannou A, Giangregorio LM. Exercise for improving outcomes after osteoporotic vertebral fracture. *Cochrane Database of Systematic Reviews* 2019, Issue 7. Art. No.: CD008618. DOI: [10.1002/14651858.CD008618.pub3](https://doi.org/10.1002/14651858.CD008618.pub3).

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Vertebral fractures are associated with increased morbidity (e.g. pain, reduced quality of life) and mortality. Therapeutic exercise is a non-pharmacological conservative treatment that is often recommended for patients with vertebral fractures to reduce pain and restore functional movement. This is an update of a Cochrane Review first published in 2013.

Objectives

To assess the effects (benefits and harms) of exercise intervention of four weeks or greater (alone or as part of a physical therapy intervention) versus non-exercise/non-active physical therapy intervention, no intervention or placebo among adults with a history of vertebral fractures on incident fragility fractures of the hip, vertebra or other sites. Our secondary objectives were to evaluate the effects of exercise on the following outcomes: falls, pain, physical performance, health-related quality of life (disease-specific and generic), and adverse events.

Search methods

We searched the following databases until November 2017: the Cochrane Library (Issue 11 of 12), MEDLINE (from 2005), Embase (from 1988), CINAHL (Cumulative Index to Nursing and Allied Health Literature, from 1982), AMED (from 1985), and PEDro (Physiotherapy Evidence Database, from 1929). Ongoing/recently completed trials were identified by searching the World Health Organization International Clinical Trials Registry Platform and ClinicalTrials.gov. Conference proceedings were searched via ISI and SCOPUS, and targeted searches of proceedings of the American Congress of Rehabilitation Medicine and American Society for Bone and Mineral Research. Search terms or MeSH headings included terms such as vertebral fracture AND exercise OR physical therapy. For this update, the search results were limited from 2011 onward.

Selection criteria

We included all randomized controlled trials and quasi-randomized trials comparing exercise or active physical therapy interventions with placebo/non-exercise/non-active physical therapy interventions or no intervention implemented in individuals with a history of vertebral fracture.

Data collection and analysis

Two review authors independently selected trials and extracted data using a pre-tested data extraction form. Disagreements were resolved by consensus, or third-party adjudication. We used Cochrane's tool for assessing risk of bias to evaluate each study. Studies were grouped according to duration of follow-up (i.e. a) 4-12 weeks; b) 16-24 weeks; c) 52 weeks); a study could be represented in more than one group depending on the number of follow-up assessments. For dichotomous data, we reported risk ratios (RR) and corresponding 95% confidence intervals (95% CI). For continuous data, we reported mean differences (MD) of the change from baseline and 95% CI. Data were pooled for Timed Up and Go test, self-reported physical function measured by the QUALEFFO-41 physical function subscale score (scale of zero to 100; lower scores indicate better self-reported physical function), and disease-specific quality of life measured by the QUALEFFO-41 total score (scale of zero to 100; lower scores indicate better quality of life) at 12 weeks using a fixed-effect model.

Main results

Nine trials (n = 749, 68 male participants; two new trials in this review update) were included. Substantial variability across the trials prevented any meaningful pooling of data for most outcomes. Risk of bias across all studies was variable; low risk across most domains in four studies, and unclear/high risk in most domains for five studies. Performance bias and blinding of subjective outcome assessment were almost all high risk of bias.

One trial reported no between-group difference in favor of the effect of exercise on incident fragility fractures after 52 weeks (RR 0.54, 95% CI 0.17 to 1.71; very low-quality evidence with control: 184 per 1000 and exercise: 100 per 1000, 95% CI 31 to 315; absolute difference: 8%, 95% CI 2 to 30). One trial reported no between-group difference in favor of the effect of exercise on incident falls after 52 weeks (RR 1.06, 95% CI 0.53 to 2.10; very low-quality evidence with control: 262 per 1000 and exercise: 277 per 1000; 95% CI 139 to 550; absolute difference: 2%, 95% CI -12 to 29). These findings should be interpreted with caution because of the very serious risk of bias in these studies and the small sample sizes resulting in imprecise estimates.

We are uncertain that exercise could improve pain, self-reported physical function, and disease-specific quality of life, because certain studies showed no evidence of clinically important differences for these outcomes. Pooled analyses revealed a small between-group difference in favor of exercise for Timed Up and Go (MD -1.13 seconds, 95% CI -1.85 to -0.42; studies = 2), which did not change following a sensitivity analysis (MD -1.09 seconds, 95% CI -1.78 to -0.40; studies = 3; moderate-quality evidence). Exercise improved QUALEFFO-41 physical function score (MD -2.84 points, 95% CI -5.57 to -0.11; studies = 2; very low-quality evidence) and QUALEFFO-41 total score (MD -3.24 points, 95% CI -6.05 to -0.43; studies = 2; very low-quality evidence), yet it is unlikely that we observed any clinically important differences. Three trials reported four adverse events related to the exercise intervention (costal cartilage fracture, rib fracture, knee pain, irritation to tape, very low-quality evidence).

Authors' conclusions

In conclusion, we do not have sufficient evidence to determine the effects of exercise on incident fractures, falls or adverse events. Our updated review found moderate-quality evidence that exercise probably improves physical performance, specifically Timed Up and Go test, in individuals with vertebral fracture (downgraded due to study limitations). However, a one-second improvement in Timed Up and Go is not a clinically important improvement. Although individual trials did report benefits for some pain and disease-specific quality of life outcomes, the findings do not represent clinically meaningful improvements and should be interpreted with caution given the very low-quality evidence due to inconsistent findings, study limitations and imprecise estimates. The small number of trials and variability across trials limited our ability to pool outcomes or make conclusions. Evidence regarding the effects of exercise after vertebral fracture in men is scarce. A high-quality randomized trial is needed to inform safety and effectiveness of exercise to lower incidence of fracture and falls and to improve patient-centered outcomes (pain, function) for individuals with vertebral fractures (minimal sample size required is approximately 2500 untreated participants or 4400 participants if taking anti-osteoporosis therapy).

PLAIN LANGUAGE SUMMARY

Exercise for improving outcomes after osteoporotic spine fracture

Researchers in Cochrane conducted an update of a 2013 Cochrane Review of the effects of exercise for people with osteoporotic spine fractures. After searching for all relevant studies up to November 2017, they found nine studies with a total of 749 people; of which two were new studies.

What are osteoporotic spine fractures and what is exercise?

Bone is a living part of your body. Throughout your life, old bone is removed and replaced with new, stronger bone. In someone with osteoporosis, old bone is removed faster than the new bone can replace it, making bones weaker and more likely to break. Exercise is often recommended for people with osteoporosis. Exercise programs may need to be modified for individuals at high risk of fracture, such as individuals with spine fractures due to osteoporosis. It is possible that exercise, if not done correctly, could increase the risk of fracture.

What happens to people with osteoporotic spine fracture who exercise?

In people over 40 years of age with a spine fracture due to osteoporosis, we do not have precise information on whether new fractures or falls happen after starting an exercise program, or about side effects and complications, especially in men.

In this review update, new findings suggest that exercise probably improves physical performance in people with spine fractures. However, it is uncertain whether exercise has any effect on pain and quality of life.

We cannot tell from our results whether exercise will cause harm, but there was evidence of adverse events related to exercise (including two rib fractures). Individuals at high risk of fracture during exercise and during transitions (e.g. rolling from one's back to stomach, handling weights) are advised to take precautions to reduce risk. Many of the interventions were delivered by a physical therapist in a research facility or center, so no conclusions can be made about exercise interventions by other health professionals or in other settings.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Exercise for improving outcomes after osteoporotic vertebral fracture

Exercise for improving outcomes after osteoporotic vertebral fracture

Patient or population: individuals with osteoporotic vertebral fracture

Settings: outpatient

Intervention: exercise

Comparison: non-exercise/non-active physical therapy intervention, no intervention or placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Control	Exercise				
Fractures Follow-up: 52 weeks	184 per 1000	100 per 1000 (31 to 315)	RR 0.54 (0.17 to 1.71)	78 (1 study)	⊕⊕⊕⊕ very low ¹	During the 12-month study, 4 participants in the exercise group sustained clinical vertebral and non-vertebral fractures and 7 participants in the control group. 84 less people out of 1000 who did exercise had a fracture (Absolute difference 8%, 95% CI 2 to 30). ²
Falls Follow-up: 52 weeks	262 per 1000	277 per 1000 (139 to 550)	RR 1.06 (0.53 to 2.10)	89 (1 study)	⊕⊕⊕⊕ very low ³	During the 12-month study, 13 participants in the exercise group and 11 in the control group had fallen over, with no between-group differences (no statistics reported). 15 more people out of 1000 who did exercise had a fall (Absolute difference 2%, 95% CI -12 to 29).
Pain Scale: VAS (0 to 10), pain sub-scale of Functional Status Index (0 to 10) Higher score indicates greater pain levels	see comment	see comment		426 (5 studies)	⊕⊕⊕⊕ very low ⁴	The range of estimates (MD between change from baseline for exercise and control groups) for pain outcomes were: -0.52 points to -2.0 points (after 4 to 12 weeks, Bennell 2010 ; Malmros 1998 ; Wang 2015); -0.45 points to -0.73 points (after 16-24 weeks, Malmros 1998 ; Wang 2015); and -0.97 points to -1.28 points (after 52 weeks, Wang 2015). Narrow 95% CIs indicate a possible effect of exercise on these pain outcomes. No between-group differences were found in two studies (Gold 2004 ; Yang 2007). MCID for the VAS (0 to 10) pain scale is typically 1-point or a 15% change (Salaffi 2004). Data were not pooled because the trials were too diverse with respect to the variability in the outcome measures chosen, the



Follow-up: 4 to 52 weeks					duration of follow-up and the interventions implemented.
Physical performance: performance-based measures TUG test Follow-up: 4 to 12 weeks	The mean TUG score in the control group for the largest study was 7.9 seconds ⁵	The TUG score in the exercise group was 1.09 seconds lower (-1.78 to -0.40)	139 (3 studies)	⊕⊕⊕⊖ moderate ⁶	One additional study (n = 89) measured walking speed (Bergland 2011). There was evidence of a small effect of exercise on maximum walking speed over 20 meters after 12 weeks (Bergland 2011). MCID for the TUG test has not been established in individuals with vertebral fractures, but the TUG test MCID typically ranges from 1.4 seconds to 3.4 seconds in other populations with chronic musculoskeletal conditions (Gautschi 2017; Wright 2011).
Physical performance: self-report questionnaires Physical function subscale from the QUALEFFO-41 Scale from 0 to 100 Lower scores indicate better physical function Follow-up: 12 weeks	The mean QUALEFFO-41 physical function score in the control group in the largest study was 22.7 points ⁷	The mean QUALEFFO-41 physical function score in the exercise group was 2.84 points lower (-5.57 to -0.11)	109 (2 studies)	⊕⊕⊕⊖ very low ⁸	Data from Bergland 2011 were pooled with Bennell 2010 (intervention combined physical therapy with exercise). There was evidence of an effect of exercise on QUALEFFO-41 physical function score in Bennell 2010, yet there were no between-group differences in Bergland 2011. Four other studies (n = 343) reported physical function questionnaire data up to 52 weeks (QUALEFFO-41, OQLQ, Oswestry Disability Index). Data were not pooled because the trials were too diverse with respect to the variability in the outcome measures chosen, the duration of follow-up and the interventions implemented. MCID has not been established for the QUALEFFO-41.
Disease-specific quality of life QUALEFFO-41 total score Scale from 0 to 100 Lower scores indicate better quality of life	The mean QUALEFFO-41 total score in the control group in the largest study was 31.8 points ⁷	The mean QUALEFFO-41 total score in the exercise group was 3.24 points lower (-6.05 to -0.43)	109 (2 studies)	⊕⊕⊕⊖ very low ⁹	Two additional studies (n = 167) measured disease-specific quality of life up to 52 weeks (QUALEFFO-41). The range of estimates (MD between change from baseline for exercise and control groups) for QUALEFFO-41 total score were -2.9 points to -8.9 points after 52 weeks (Bergland 2011; Evstigneeva 2016). ¹⁰ One other study (n = 74) reported quality of life outcomes from the OQLQ. Data were not pooled because the trials were too diverse with respect to the variability in the outcome measures chosen, the duration of follow-up and the interventions implemented. MCID has not been established for the QUALEFFO-41.

Follow-up: 12 weeks						
Adverse events	see comment	see comment	Not estimable	447 (6 studies)	⊕⊕⊕⊕ very low ¹¹	There were 4 adverse events related to the intervention: costal cartilage fracture, rib fracture, knee pain, and irritation to tape.
Follow-up: 12 to 52 weeks						

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% CI) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **MCID:** minimal clinically important difference; **VAS:** visual analogue score; **MD:** mean difference; **TUG:** Timed Up and Go; **QALFFO-41:** Quality of Life Questionnaire of the European Foundation for Osteoporosis; **OQLQ:** Osteoporosis Quality of Life Questionnaire; .

GRADE Working Group grades of evidence

⊕⊕⊕⊕ **High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

⊕⊕⊕⊕ **Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

⊕⊕⊕⊕ **Low quality:** 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 quality:** We are very uncertain about the estimate.

¹Downgraded two levels for study limitations due to incomplete reporting of data and because there was no clear description of how incident fragility fractures were confirmed. Downgraded one level for imprecision due to imprecise results or sparse data. Downgraded one level due to indirectness due to the evaluation of incident fracture as a secondary outcome.

²[Evstigneeva 2016](#) did not describe how incident fragility fractures were confirmed in their study and it was not possible to separately report the symptomatic and asymptomatic fractures.

³Downgraded one level for study limitations due to incomplete reporting of data. Downgraded one level for indirectness due to the evaluation of incident falls as a secondary outcome. Downgraded one level for imprecision due to sparse data.

⁴Downgraded two levels for study limitations due to lack of blinding in all trials, inadequate/unclear random sequence generation or allocation concealment in more than 1 trial, and incomplete reporting of data in more than 1 trial. Downgraded one level for imprecision due to sparse data. Downgraded one level for inconsistency due to heterogeneity in the results.

⁵The baseline Timed Up and Go test score for the control group from [Bergland 2011](#) was used.

⁶Downgraded one level for study limitations due to lack of blinding in all trials, inadequate/unclear random sequence generation or allocation concealment in more than 1 trial, and incomplete reporting of data in 1 trial.

⁷The baseline QALFFO-41 physical function and total scores for the control group from [Bergland 2011](#) was used.

⁸Downgraded one level for study limitations including lack of blinding due to the nature of the intervention. Downgraded one level for inconsistency due to heterogeneity in the results. Downgraded one level for imprecision due to sparse data.

⁹Downgraded one level for study limitations including lack of blinding due to the nature of the intervention. Downgraded one level for inconsistency due to heterogeneity in the results. Downgraded one level for imprecision due to sparse data.

¹⁰[Evstigneeva 2016](#) reported on seven domains for QALFFO-41 (physical function subscale in three separate scores - activities of daily living, jobs around the house, and mobility).

¹¹Downgraded two levels for study limitations because of the high probability of selective outcome reporting; there was no clear description of how the adverse events were recorded or monitored in the methods of the included trials; and insufficient details to report on the distribution between exercise and control groups. Downgraded one level for imprecision due to inconsistency in findings and sparse data.

BACKGROUND

Description of the condition

Osteoporosis is a disease of the skeleton characterized by low bone mineral density and deterioration of bone tissue, resulting in an increased risk of fragility fracture (WHO 2003). A fragility fracture is a fracture that occurs with minimal trauma, such as a fall from standing height. The wrist, hip and vertebra are the most common sites for osteoporotic fragility fractures (Kanis 2001). Fragility fractures, particularly those of the hip and vertebra, are associated with increased mortality and significant morbidity including impaired physical performance, pain, deformity, sleep disturbance, depression, fear of future falling and fracture, and reduced quality of life (Adachi 2001; Cauley 2000; Papaioannou 2002; Petrella 2000; Wiktorowicz 2001).

Estimating the prevalence and impact of vertebral fractures is difficult; only about 30% of vertebral fractures come to clinical attention as they depend on a report of pain or height loss that triggers the clinician to order a radiograph (Papaioannou 2002). Further, many fractures are not reported even when present on X-ray (Papaioannou 2003b). Estimates of the prevalence of vertebral fractures among men and women have been reported to be similar; prevalent vertebral deformities were found in 23.5% of women and 21.5% of men aged 50 years and older in a Canadian population-based study (Jackson 2000). Similarly, in the USA, a prevalence rate of 25.3% in women 50 years and over has been reported (Cooper 1993). Among individuals presenting with a fragility fracture in the UK, 25% had evidence of vertebral deformity on X-ray (Gallacher 2007). The absolute risk of a subsequent vertebral fracture among women with a prevalent vertebral fracture and osteoporosis based on bone mineral density has been reported to be 50%, compared to 9% among women with no fracture and normal bone mineral density (Cauley 2007). Increasing age and a history of fragility fracture predict subsequent fractures independent of bone mineral density (Kanis 2004). Therefore, individuals who experience vertebral fracture should be targeted for fracture prevention strategies.

Description of the intervention

The management of osteoporosis is multi-modal, and includes pharmacological and non-pharmacological interventions (Avenell 2014; Wells 2008a; Wells 2008b; Wells 2008c). Exercise alone or as part of physical therapy management is often recommended as a beneficial non-pharmacological treatment to slow the rate of bone loss. National and international osteoporosis organizations emphasize the importance of physical activity for preserving bone health. Recent knowledge synthesis activities established exercise recommendations for individuals with osteoporotic vertebral fractures that encourage resistance, balance and aerobic exercise training (Giangregorio 2014a; Giangregorio 2015). Exercise interventions designed for individuals with vertebral fracture may include postural correction, challenging balance training and modified trunk and lower extremity muscle strengthening exercises in addition to moderate-intensity aerobic physical activity. The goals of such exercise interventions are to regain or maintain normal spine curvatures, increase spine stability, improve functional performance, and avoid postures and physical activities that may increase the risk of falls and fracture (Bennell 2010; Giangregorio 2014a; Howe 2011).

How the intervention might work

Meta-analyses exploring the impact of exercise on bone mineral density in postmenopausal women suggest that weight-bearing exercise or resistance training may decrease the rate of bone loss in women, but the effect of bone mineral density measured by dual energy X-ray absorptiometry (DXA) may vary with exercise mode (Kelley 1998; Martyn-St James 2009; Moayyeri 2008). Studies to date include primarily healthy women without a history of vertebral fracture or in some cases with normal bone mineral density, which may not be generalizable to individuals with vertebral fractures. Exercise may have important effects on bone strength that are not reflected in DXA-based assessments of bone mineral density (Polidoulis 2012). There is also evidence that exercise improves muscle strength and balance and prevents falls (Sherrington 2011), which may indirectly prevent fractures. Moreover, exercise can improve physical performance and prevent functional impairment in older adults (Chou 2012; Pahor 2014). However, as with studies of effects of exercise on bone mineral density, restrictive study sampling limits generalizability to individuals with vertebral fractures. Further, individuals with vertebral fractures may have hyperkyphosis (excessive curvature of the thoracic spine), which increases spinal loading (Briggs 2007; Bruno 2012). Exercises that are aimed at reducing kyphotic posture may reduce the risk of vertebral fracture.

Why it is important to do this review

Evidence-based clinical practice guidelines for exercise prescription specific to individuals with vertebral fractures are lacking. Exercise guidelines developed for healthy older adults may not be appropriate for individuals with vertebral fractures; the types of exercises and their intensity may need to be modified. Systematic reviews on the topics of interventions for improving physical performance after hip fracture, and rehabilitation after distal radius fracture have been developed by Cochrane (Handoll 2015; Handoll 2011). A systematic review of exercise interventions in individuals with vertebral fracture, published in 2010, highlighted the paucity of work in the area (Dusdal 2010). Several of the trials in the latter review did not include exclusively individuals with vertebral fracture or were not randomized trials. The first edition of this Cochrane Review reported that only seven trials evaluated the efficacy of exercise after vertebral fracture with notable limitations, such as bias, small sample sizes, and lack of long-term follow-up (Giangregorio 2013). Newer trials may have been published since the original search was performed. A recent international consensus process on research priorities for this field revealed that a large-scale randomized controlled trial in people with osteoporotic vertebral fracture is needed to determine whether exercise does more good than harm (Giangregorio 2014b). It is possible that the risk of fracture might be increased with exercise in individuals with vertebral fractures, but the risk has not been quantified. Therefore, we conducted an updated synthesis of trials on the effect of exercise interventions for improving outcomes for individuals with vertebral fracture with the goal of providing a comprehensive and current review on knowledge gaps and best evidence to inform the design of future randomized controlled trials.

OBJECTIVES

The primary objective of this review update was to evaluate the benefits, i.e. lower incidence of fractures, and harms, i.e. adverse

events, of exercise interventions of four weeks or greater (alone or as part of a physical therapy intervention) versus non-exercise/non-active physical therapy intervention, no intervention or placebo among adults with a history of osteoporotic vertebral fracture(s).

The secondary objectives were to evaluate the effect of exercise interventions of four weeks or greater (alone or as part of a physical therapy intervention) versus non-exercise/non-active physical therapy intervention, no intervention or placebo on the following health-related outcomes among adults with a history of osteoporotic vertebral fracture(s): incident falls; pain; physical performance; self-reported physical function; health-related quality of life; posture; muscle function; balance; bone mineral density of the lumbar spine or hip measured using DXA; fear of falling; and patient global assessment of success. We also described adherence to the interventions.

METHODS

Criteria for considering studies for this review

Types of studies

We considered for inclusion all randomized controlled trials (including those in which the treatment allocation was inadequately concealed) or quasi-randomized trials comparing an exercise intervention (alone or as part of a physical therapy intervention) with a non-exercise/non-active physical therapy intervention, no intervention or placebo implemented in individuals with a history of vertebral fracture.

Types of participants

We included studies of men and women over the age of 40 years with a history of non-traumatic or minimal trauma osteoporotic fracture of one or more vertebrae. A non-traumatic fracture was defined as a fracture that occurs spontaneously. A minimal trauma fracture was defined as a fracture that occurs following a:

- fall from standing height;
- fall from sitting position;
- fall from supine position (bed or reclining deck chair < 1 meter high);
- fall after having missed one to three steps in a staircase;
- movement outside of the typical plane of motion or coughing ([Bessette 2008](#)).

Types of interventions

Treatment: trials that involved exercise of any kind, such as: muscle strengthening or resistance training exercises, aerobic exercise, balance training, Tai Chi, or individualized exercise prescribed by a physical therapist were included. Trials examining modalities or devices that did not include an active physical activity component were excluded. Trials that included co-interventions were not excluded. For example, multi-modal physical therapy interventions were included if one group received exercise as part of the multi-modal intervention and the comparison group received a non-exercise intervention or no intervention. Trials had to include an intervention of at least four weeks duration with subsequent outcome assessment; studies with interventions of less than four weeks duration or with all outcome assessments prior to four weeks of intervention were excluded. All variations of frequency, intensity and duration of intervention during each session were considered.

Comparators were: a) non-exercise/non-active physical therapy intervention (e.g. educational intervention); b) no intervention; or c) placebo.

Types of outcome measures

Major outcomes

1. Incident fragility fractures of the hip, vertebra or other sites:
 - symptomatic or asymptomatic fragility fractures confirmed on X-ray; or
 - reduction in vertebral height of greater than 20% ([Schousboe 2008](#)), as measured using the Vertebral Fracture Assessment protocol with DXA.
2. Incident falls:
 - self-reported falls;
 - falls documented in medical records.
3. Pain:
 - self-report questionnaires specifically developed to assess pain.
4. Physical performance:
 - performance-based measures of physical performance, e.g. documented use of walking aids, six-minute walk test or Timed Up and Go test;
 - self-reported questionnaires specifically developed and validated to assess physical function; subscales from validated self-report questionnaires related to physical function, e.g. SF-36
 - other indices of physical performance as described in each study.
5. Health-related quality of life:
 - disease-specific self-report questionnaires;
 - generic self-report questionnaires.

Disease-specific quality of life questionnaires have a better ability to clinically characterize quality of life specific to osteoporosis and to assess changes in quality of life over time or treatment in individuals with osteoporotic vertebral fractures ([Lips 2005](#); [Oleksik 2000](#)). Generic health-related quality of life tools are designed to measure quality of life across a broad spectrum of disease and disability. For clinical trials reporting on both disease-specific and generic health-related quality of life, we planned to use the disease-specific quality of life measure for primary interpretation of the effects of exercise.
6. Adverse events (other than fragility fractures or falls):
 - serious adverse events that may or may not be related to the study intervention, defined as “any untoward medical occurrence that results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity” ([FDA 1995](#));
 - adverse events, defined as “any unfavourable and unintended sign, symptom, or disease” ([FDA 1995](#)). The adverse event should be temporally associated with study participation.

Minor outcomes

7. Posture: measures of postural alignment or spine curvature, e.g. measurements made using clinical devices such as an inclinometer, or from radiographs.

8. Muscle Function

- Muscle strength or endurance of key muscle groups, e.g. back extensors or lower limb muscles, measured quantitatively.

9. Balance

- Balance, including self-report questionnaires and performance-based measures e.g. assessments of balance performance using a force plate.

10. Bone mineral density of the lumbar spine or hip measured using DXA.

11. Fear of falling: self-report questionnaires developed to assess fear of falling, e.g. Fall Self-Efficacy Scale- International.

12. Patient global assessment of success, e.g. self-report rating scales to assess global health or disease severity.

Search methods for identification of studies

Electronic searches

We searched the following databases: the Cochrane Library (November 2019, Issue 3), MEDLINE (2005 to 20 March 2019), Embase (1988 to 20 March 2019), CINAHL (Cumulative Index to Nursing and Allied Health Literature, 1982 to 2 November 2017), AMED (1985 to 2 November 2017), and PEDro (Physiotherapy Evidence Database, www.pedro.fhs.usyd.edu.au/index.html, 1929 to 2 November 2017). Ongoing and recently completed trials were identified by searching the World Health Organization International Clinical Trials Registry Platform and ClinicalTrials.gov (to 20 March 2019). We did not apply any language restrictions. MEDLINE searches were undertaken using MeSH headings and text words for vertebral fracture, exercise and physical therapy. The MEDLINE strategy ([Appendix 1](#)) was modified for use in the Cochrane Library, Embase, CINAHL, PEDro, World Health Organization International Clinical Trials Registry Platform, and ClinicalTrials.gov ([Appendix 2](#); [Appendix 3](#); [Appendix 4](#); [Appendix 5](#); [Appendix 6](#); [Appendix 7](#); [Appendix 8](#)). For this update, the search results were limited from November 2011 onward. The search update process was run in three stages: the first search was run in November 2015; a second top-up search was run in November 2017, and a third on March 2019.

Searching other resources

We searched the reference lists of included articles for additional references. We searched conference proceedings using ISI and SCOPUS. We also searched the conference proceedings available online for the American Congress of Rehabilitation Medicine and American Society for Bone and Mineral Research.

Data collection and analysis

Selection of studies

Two review authors (LG, JG, or MP) reviewed the title, abstract and descriptors of identified studies for possible inclusion. From the

full text, two review authors (LG, NM, JG, MP, or JT) independently assessed potentially eligible trials for inclusion. The percentage agreement between review authors during level two screening was 94%. Any disagreement was resolved through discussion. We contacted authors of articles when additional information was needed.

Data extraction and management

Two review authors independently reviewed each trial (LG, NM, JG, MP, or JT), and extracted data using a pre-tested data extraction form. The extraction form was tested on two articles by the two review authors. Any disagreement was resolved by consensus or third party adjudication. Review authors did not review their own trials.

Assessment of risk of bias in included studies

Two review authors independently assessed risk of bias (LG, NM, JG, or JT). An assessment tool ([Table 1](#)) was developed based on the recommendations in chapter 8 of the *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2011](#)). The following criteria were used to evaluate risk of bias: random sequence generation; allocation concealment; blinding of participants and personnel; blinding of outcome assessors; addressing incomplete outcome data; and selective outcome reporting. We reviewed each study for the presence or absence of each criterion, and coded it as low risk of bias, unclear risk of bias (uncertain risk of bias) or high risk of bias. Any disagreement regarding risk of bias was resolved by consensus.

Measures of treatment effect

We calculated mean differences (MD) (95% confidence intervals, (CI)) for continuous outcomes. We calculated risk ratios (RR) and corresponding 95% CIs for binary outcomes with available data (fractures, falls). We planned to use the standardized mean difference (SMD), as described in chapter 9.4.6 of the *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2011](#)), to pool an outcome measured with different scales but the data available for pooling did not permit these calculations. We were only able to pool data for three outcomes because of the limited number of trials. To convert CI to standard deviations, we used the methods described in chapter 7.7.3.2 of the *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2011](#)).

Unit of analysis issues

No trials were included in pooling that required adjustment for clustering, or correction for design effect in clustered trials that did not adjust for clustering in the analyses and therefore, we did not need to impute any intra-cluster correlation (ICC) to calculate the Variance Inflation Factor (VIF) as described in our protocol.

Dealing with missing data

We reported levels of attrition for included trials. We contacted the authors of trials to provide missing data when data were not reported for some outcomes or groups. We planned to impute missing data and test the effects of imputation on the conclusions that we made ([Little 1987](#)).

We analyzed participant data in the group to which they were allocated, independent of whether they received the allocated intervention or not. We planned to re-analyze data for trials where

participants were not analyzed in the group to which they were allocated, if there was sufficient information in the trial report or if the data could be retrieved from the study authors. One study (Malmros 1999) did not perform an intention-to-treat analysis, so the raw data were obtained and missing data were imputed using the last observation carried forward method, with the exception of one subscale of the balance test for one individual, where we imputed the mean of all participants in that group because no data were available for any time point. We performed an analysis of variance on the change from baseline for all available time points at which outcomes were assessed (one or more outcomes were assessed at five weeks, 10 weeks and 22 weeks).

Assessment of heterogeneity

We used the χ^2 test and I^2 statistic to quantify any unexplained heterogeneity, where an I^2 of less than 25% was considered low heterogeneity, an I^2 of 25 to 50% was considered moderate heterogeneity and an I^2 of over 50% was considered high heterogeneity (Higgins 2003).

Assessment of reporting biases

We planned to use funnel plots to evaluate publication bias, but there were not enough trials to do so. We had planned to use the capture-mark-recapture technique (e.g. the Horizon Estimate) to estimate the total number of articles in the domain of clinical trials of exercise for individuals with vertebral fracture and the proportion of these articles we were able to capture with our search strategy (Kastner 2007; Kastner 2009). We could not perform the Horizon Estimate analysis due to the small number of studies.

Data synthesis

We planned to pool results of exercise with comparable outcomes using a random-effects approach (95% CI) in Review Manager (RevMan 5). In almost all cases, we did not pool data because of the heterogeneity across trials and the number of trials, or the lack of available data in publications and after queries to authors. In the instance of only two available trials with comparable outcomes for each of Timed Up and Go and Quality of Life Questionnaire of the European Foundation for Osteoporosis (QALEFFO-41) total score and physical function subscale score, we used a fixed-effect approach (95% CI) in RevMan 5. A sensitivity analysis was conducted to compare the results from a meta-analysis of Timed Up and Go test performance with and without Bennell 2010 (excluded in the first edition of this review because there was no comparison group that received the same intervention with the exception of exercise) (Giangregorio 2013).

The comparative risks, relative effects, number of participants studied and grade of evidence related to the following outcomes are presented in a 'Summary of findings' table as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011): fractures, adverse events, falls, pain, physical performance

(specifically Timed Up and Go and walking speed), and health-related quality of life. The GRADE approach was used to rate the quality of the body of evidence for each outcome (Higgins 2011).

Subgroup analysis and investigation of heterogeneity

Heterogeneity between trials may be related to a number of participant-related or intervention-related variables. We had planned to test a number of hypotheses regarding sources of heterogeneity, if applicable, in subgroup analyses, where subgroups would be defined according to the following: level of supervision (continual, intermittent, none), site (home-based versus center-based), type of exercise, target goal of exercise or positioning during exercise (e.g. strength training, aerobic training, postural retraining, combination, open versus closed chain, dynamic versus static), dose (frequency and duration), intensity, compliance (greater than 75% of sessions versus less than or equal to 75% of sessions), randomized versus quasi-randomized and individual versus cluster-randomized designs. We did not perform these *a priori* subgroup analyses because of the limited number of trials and inability to pool data, but retain the description of our plan to inform future work. We planned to use meta-regression to explain the heterogeneity attributable to these variables. We hypothesized that participants who were continually supervised (i.e. supervision during every session) would achieve greater benefit versus intermittent supervision (i.e. supervision during some, but not all sessions), which would be better than no supervision (i.e. preliminary instructions provided but no consistent monitoring, or less than 5% of sessions monitored) because a greater degree of supervision would provide more opportunity for individualized prescription, feedback regarding proper form, and less fear of adverse events.

Sensitivity analysis

We had planned to use sensitivity analyses to explore the effect of study quality (adequate versus inadequate allocation concealment and blinding) and number of participants with multiple spine fractures at baseline by including them as potential predictors in the meta-regression when testing for heterogeneity. We were not able to do this because of the limited number of trials.

RESULTS

Description of studies

We identified 1772 references (Figure 1). After reviewing the titles and abstracts, 41 full-text articles, 15 abstracts, and one clinical trials record were retrieved in the case of journal publications. When abstracts were identified from conference proceedings, they were cross-referenced with the results of the search to ensure that data were not included twice. At least three attempts were made via email to contact the authors of the abstracts to provide additional information when necessary.

Figure 1. PRISMA study flow diagram.

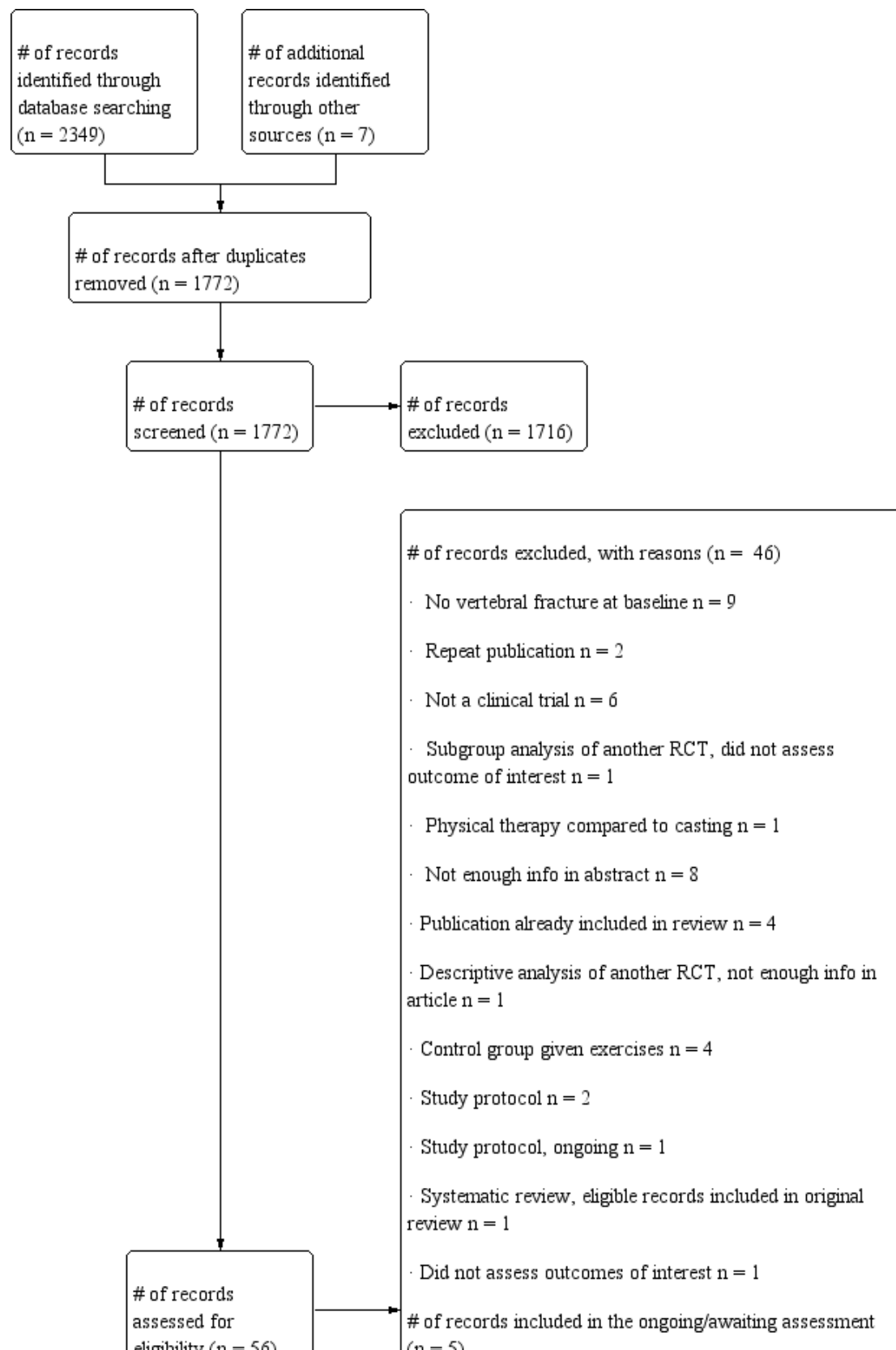
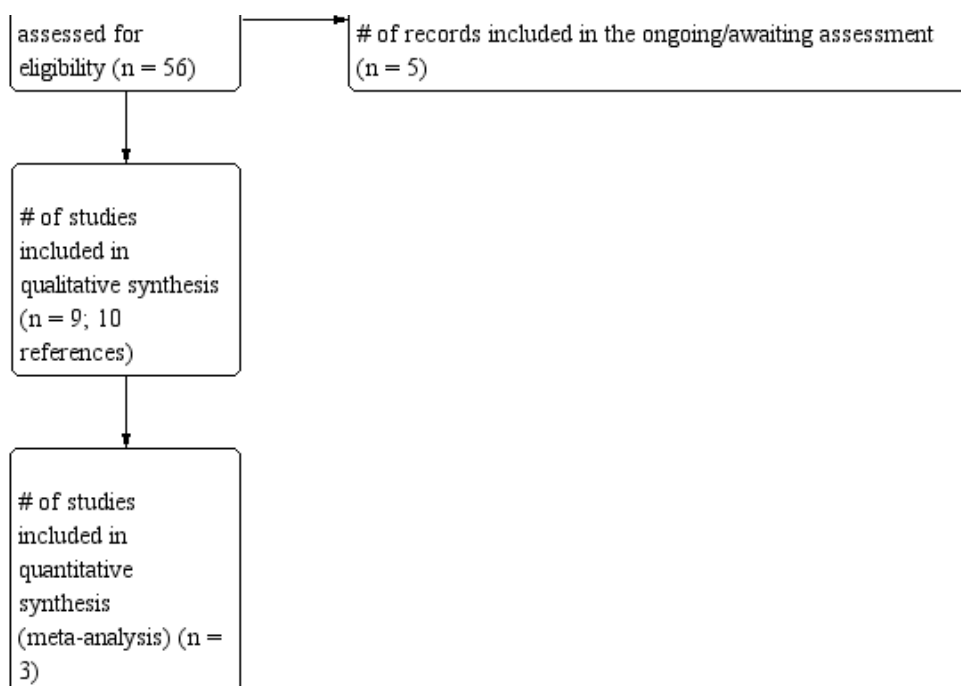


Figure 1. (Continued)



Nine studies from 10 full-text journal articles (Bennell 2010; Bergland 2011; Bergstrom 2011; Gold 2004; Malmros 1998; Olsen 2014; Papaioannou 2003a; Yang 2007; Wang 2015; Evstigneeva 2016) with a total of 749 participants were eligible for inclusion (including two new trials; participant = 68 males). One study (Gold 2004) was a cluster-randomized trial, where each cluster was a retirement residence. The variance inflation factor (used to adjust the standard deviation in the treatment and control groups which accounts for average cluster size and the ICC) was not reported for this trial, but the site was controlled for in the analysis. One full-text article (Olsen 2014) was a secondary analysis of Bergland 2011. Therefore, we only extracted the new data on the effect of exercise on self-reported number of falls and fear of falling and did not report as a separate study or assess for risk of bias.

Bennell 2010 and Wang 2015 were the only trials to include men. In Bennell 2010, four of the 11 participants in the intervention group and none of the nine in the control group were men. In Wang 2015, 24 of the 46 participants in the intervention group and nine and 11 participants in the two control groups were men. Eight articles were published in English, and two articles were published in Chinese and translated to English (Yang 2007; Wang 2015). Six published abstracts were identified (Miyakoshi 2010; Evstigneeva 2016; Evstigneyeva 2013; Evstigneeva 2013; Mihailov 2011; Nemes 2014), however, attempts to contact the authors to confirm unreported details or data were mostly unsuccessful and three abstracts were excluded. Three abstracts were secondary analyses of the study by Evstigneeva 2016 (Evstigneeva 2016; Evstigneyeva 2013; Evstigneeva 2013). Bergstrom 2016, Giangregorio 2017, and Pompa 2016 were published abstracts, Giangregorio 2018 was a publication in press, and Stanghelle 2018 was an ongoing trial identified as awaiting assessment, which may be included in a future update of our review. All interventions are described in detail in [Characteristics of included studies](#). Characteristics that may introduce heterogeneity in the findings are described below, and

include participant characteristics, exercise frequency, intensity and duration, the setting and level of supervision, adherence to the intervention, the outcomes chosen, as well as co-interventions or comparator group activities.

Participant characteristics

The nine trials included 749 participants of which 68 were male. All trials had history of vertebral fracture as an inclusion criterion, but the way vertebral fracture was defined varied between studies. Studies defined fracture based on the presence of symptoms (Malmros 1998), or on morphometric changes observed on X-ray, and among those, the definition varied e.g. no definition provided (Bergland 2011; Bergstrom 2011; Evstigneeva 2016; Wang 2015), height reduction of $\geq 15\%$ (Papaioannou 2003a) or $\geq 20\%$ (Bennell 2010; Gold 2004; Yang 2007), using DXA or radiography. Four studies (Bennell 2010; Evstigneeva 2016; Malmros 1998; Wang 2015) had pain as an inclusion criterion; future studies that aim to evaluate effects on pain associated with vertebral fracture should consider studying effects on individuals with pain at baseline. Variability within a study population or between populations in the severity of vertebral fracture or in the presence of symptoms may result in variable effects of exercise or in generalizability.

Frequency, intensity and duration of the exercise intervention, and duration of follow-up

There was considerable diversity in the frequency, intensity, and duration of interventions, as well as follow-up periods. Bergland 2011, Bergstrom 2011, Evstigneeva 2016, and Malmros 1998 reported a recommended exercise frequency of two times per week; Bennell 2010; Gold 2004; Papaioannou 2003a, and Yang 2007 reported a frequency of three times per week; and Wang 2015 reported a recommended daily exercise frequency. Bennell 2010 reported that muscle strengthening exercises were to be performed three times a week, and that posture training and range of motion

exercises should be performed daily. None of the studies reported a specific intensity; when intensity was reported it was often adjusted according to clinical presentation. Four of the studies evaluated exercise interventions that were performed for one to three months (Bennell 2010; Bergland 2011; Malmros 1998; Yang 2007). Bergstrom 2011 implemented a four-month intervention where the stated goal was back muscle strengthening; 30 repetitions were prescribed for each exercise. Gold 2004 implemented a six-month exercise intervention (a thrice-weekly group exercise class). Papaioannou 2003a asked participants to exercise at home thrice weekly for 12 months. Evstigneeva 2016 asked participants to exercise at a center twice weekly for 12 months, and if unable to regularly attend exercise sessions at the center, participants were allowed to perform exercises at home. Wang 2015 asked participants to perform low back strengthening exercises at home daily for 12 months. In all studies, outcome assessment occurred before and immediately after the intervention period. Two studies had a follow-up outcome assessment when the exercise ceased and also some time later: Bergland 2011 implemented a twice-weekly exercise class for 12 weeks, and outcome assessment was performed at 12 weeks and 12 months, and Malmros 1998 implemented a twice-weekly exercise class for 10 weeks, and outcome assessment was performed at 10 weeks and 22 weeks. Papaioannou 2003a performed an interim analysis at the half way point, six months after randomization. Wang 2015 performed outcome assessment once prior to percutaneous kyphoplasty surgery, and at three days, two weeks, one, six, and 12 months post-surgery. Gold 2004 also collected interim data at the half-way point, three months after randomization, yet, the data were not reported, and were not available for inclusion in this review. Gold 2004 had a Phase two where the control group received the intervention for six months, and the intervention group practiced self-maintenance. We did not include the outcomes of Phase two as part of this review as it was no longer a randomized controlled trial.

Setting and supervision

Two studies included home-based exercise (Papaioannou 2003b; Wang 2015), two studies combined a clinic-based intervention that included a home exercise program (Bennell 2010; Malmros 1999), and the remaining five studies implemented a center-based intervention (Bergland 2011; Bergstrom 2011; Evstigneeva 2016; Gold 2004; Yang 2007). Evstigneeva 2016 allowed their participants to perform the structured exercise program at home guided by a booklet and DVD discs (without supervision) if regular attendance of the center-based sessions was impossible and proper instruction on the exercises was provided. Five studies were supervised by a physical therapist (Bennell 2010; Bergland 2011; Bergstrom 2011; Gold 2004; Malmros 1998), whereas, one study was supervised by a kinesiologist (Papaioannou 2003a). Three studies did not report on the type or training level of the instructor that led the exercise intervention (Evstigneeva 2016; Wang 2015; Yang 2007). Four studies reported continual supervision (Bergland 2011; Bergstrom 2011; Evstigneeva 2016; Gold 2004), three studies reported intermittent supervision (Bennell 2010; Malmros 1998; Papaioannou 2003a), and for two studies, the level of supervision was unclear (Yang 2007; Wang 2015). Four interventions were conducted for six months or longer using the following arrangements: 1) intermittent supervision was provided for the first six months followed by telephone follow-up for the final six months (Papaioannou 2003a); 2) continual supervision was provided throughout the six-month intervention in a group

exercise class (Gold 2004); 3) continual supervision was provided throughout the one-year intervention in a one-on-one session with an instructor at a center or without supervision at home if unable to attend center-based classes (Evstigneeva 2016); and 4) the level of supervision was unclear (Wang 2015).

Adherence

The four studies of exercise interventions of four to 12 weeks (Bennell 2010; Bergland 2011; Malmros 1998; Yang 2007) can provide insight into short-term adherence to exercise in individuals with vertebral fracture. The way adherence was reported varied: Malmros 1998 reported an average adherence of 100% (90% to 100%); Bennell 2010 reported that eight of 11 (73%) of participants achieved 100% adherence to therapy sessions and a median adherence to home exercise sessions of 95%, with an overall minimum and maximum adherence of 34% and 100%; Bergland 2011 reported that the mean number of completed sessions was 19.5 out of a possible 24 (standard deviation (SD) 4.4), and that 24% of participants completed < 19 sessions; and Yang 2007 reported that all participants completed 100% of treatment. Regarding longer-term adherence, Gold 2004 reported a mean attendance to exercise classes of 58% over six months. In Papaioannou 2003a, 62% of participants reported performing their exercises at home three times per week at six months, which declined to 46% at 12 months. In Evstigneeva 2016, 89.2% of participants missed less than 20 sessions over the 12 months; whereas, four participants reported missing more than 20 sessions. An average adherence was not reported in Bergstrom 2011, but it was noted that two individuals could not complete the intervention, two individuals in the intervention group were lost to follow-up and 28 individuals achieved 90% attendance to exercise sessions. In Wang 2015, adherence was not specifically reported; yet, it was mentioned that the participant compliance to the intervention was relatively poor. For six studies (Bergland 2011; Bergstrom 2011; Evstigneeva 2016; Gold 2004; Malmros 1998; Papaioannou 2003a), it was not clear whether dropouts were considered in the estimates of adherence.

Major outcomes

Below we report the outcomes and associated measures that were evaluated in the included studies. Webber 2003 included a comparison of vertebral height as a continuous measure obtained in a subset of participants from Papaioannou 2003a, but this study was excluded for several reasons (see Characteristics of Excluded studies). Since Olsen 2014 was a secondary analysis of the same sample as Bergland 2011, only new data on the effect of exercise on self-reported number of falls and fear of falling were extracted, which were not previously published by Bergland 2011. Evstigneeva 2016, Evstigneeva 2013 and Evstigneeva 2013 reported on a subset of outcomes (Quality of Life Questionnaire of the European Foundation for Osteoporosis (QUALEFFO-41), Timed Up and Go test, and stabilimetry- Sit-to-Stand Weight Transfer, Body Weight Rising Index, and Tandem Walk and Sway) from a published full-text article by Evstigneeva 2016.

Incident fragility fractures of the hip, vertebral and other sites

One study evaluated fractures as a secondary outcome (Evstigneeva 2016). Evstigneeva 2016 did not describe how incident fragility fractures were confirmed in their study and they did not report whether they included symptomatic or asymptomatic fractures or both. Fractures were reported as adverse events in some studies (Table 2).

Incident falls

One study evaluated the self-reported number of falls in the intervention and control groups during a 12-month follow-up period as a secondary analysis (Olsen 2014). Falls were reported as adverse events in some studies (Table 2).

Pain

Pain was measured in six studies (Bennell 2010; Evstigneeva 2016; Gold 2004; Malmros 1998; Yang 2007; Wang 2015): Bennell 2010 used an 11-point scale to assess pain intensity on movement in the previous week and pain intensity at rest in the previous week; Gold 2004 measured pain with activities using the pain subscale of the Functional Status Index; Malmros 1999 used an 11-point scale to assess pain intensity in the previous week; and Yang 2007 used a 10 cm visual analogue scale to assess pain intensity (0 = no pain, 10 cm = worst pain, no reference timeframe provided), but, the means provided in the results are all higher than 10, suggesting that they converted the data to a 100 mm scale. Malmros 1999 also used a five-point categorical scale for participants to rate analgesic use, but we did not analyze these data. Evstigneeva 2016 measured pain using a 100 mm visual analogue scale, but they did not specify any further details on the scale and only reported baseline data. Wang 2015 did not specify details on the scale or reference timeframe they used for their assessment of pain, and the means provided in the results are all less than 10, suggesting a 10 cm scale.

Physical performance: performance-based measures

Several studies examined between-group differences in physical performance; walking speed (Bergland 2011), the Timed Up and Go test (Bennell 2010; Bergland 2011; Evstigneeva 2016; Yang 2007; Papaioannou 2003a), and the functional reach test (Bergland 2011). Yang 2007 measured time to get up from a supine position.

Physical performance: self-reported physical function questionnaires

Several studies examined between-group differences in self-reported physical function using subscales of a quality of life tool: the QUALEFFO-41 (Bennell 2010; Bergland 2011 - physical function subscale; Evstigneeva 2016- three physical function subscales-activities of daily living, jobs around the house, mobility) or the Osteoporosis Quality of Life Questionnaire (OQLQ) (Papaioannou 2003a). Physical function was measured using a modified or regular version of the self-reported Oswestry Disability Index (Malmros 1998; Wang 2015). A 10-point scale to assess restriction of everyday activities in the previous week was used in Bennell 2010.

Health-related quality of life (disease-specific and generic)

Quality of life was measured using disease-specific (Bennell 2010; Bergland 2011; Evstigneeva 2016; Papaioannou 2003a) questionnaires (QUALEFFO-41 and Osteoporosis Quality of Life Questionnaire (OQLQ)) and generic (Bennell 2010; Bergland 2011; Papaioannou 2003a) questionnaires (Assessment of Quality of Life (AQoL), the General Health Questionnaire, and the Sickness Impact Profile). Malmros 1999 used a non-validated questionnaire that asked participants to score their quality of life compared to their perceived quality of life at baseline; we did not analyze these data. QUALEFFO-41 and OQLQ were the only two osteoporosis-specific quality of life questionnaires included in the review. Although there were several studies that measured quality of life (QOL), we chose not to pool these data (except for QUALEFFO-41) because of the

diversity in follow-up times and lack of required data or composite scores for pooling.

Adverse events

No studies evaluated between-group differences in adverse events (Table 2). Also, no studies specifically indicated that adverse events were included as an outcome, or described a method for assessing and recording adverse events and their severity throughout the trial.

Minor outcomes

Posture

Posture was measured in Bennell 2010 with an inclinometer. Bergstrom 2011 indicated that sagittal thoracic spine curvature in maximum extension was measured using a kyphometer. Evstigneeva 2016 measured occiput-to-wall distance as an indicator of the severity of thoracic kyphosis.

Muscle function

Trunk muscle endurance was measured by Bennell 2010 using the Timed Loaded Standing test. Bergstrom 2011 measured back extensor muscle strength in standing using an isometric dynamometer. Back extensor muscle strength was measured by Malmros 1998 with a strain gauge. Gold 2004 measured peak isometric torque of the back extensor muscles using the B-200 Isostation.

Balance

Balance was measured with center of pressure variability recorded using a force plate (Malmros 1998; Papaioannou 2003a). Evstigneeva 2016 assessed balance and its dynamics using stabilimetry with a computed posturographic system for three tests: 1) weight bearing/squat (percentage of body weight); 2) sit-to-stand- weight transfer (in seconds) and left/right weight symmetry (percentage of body weight); and 3) tandem walk and sway (degrees per second).

Bone mineral density of the lumbar spine or hip measured using DXA

Papaioannou 2003a and Wang 2015 were the only studies to measure bone mineral density.

Fear of falling

Olsen 2014 was the only study to measure fear of falling using the Falls Efficacy Scale-International (FES-I).

Patient assessment of global success

None of the included studies measured patient assessment of global success.

Co-Interventions and comparison groups

Bennell 2010 implemented a physical therapy intervention that included manual therapy, massage and taping in addition to a home exercise program. The control group did not receive any intervention. An exercise intervention implemented by Gold 2004 was combined with a twice-weekly coping class for 45 minutes designed to reduce psychological concerns common to individuals with vertebral fracture. The control group attended a once-weekly 45-minute class where general health concerns

were discussed. Wang 2015 examined two treatment plans (low back muscle strengthening exercises versus salmon calcitonin-50 IU (International Units) daily doses through intramuscular injection for 2 weeks followed by 50 IU once every other day for another two weeks) of individuals who underwent percutaneous kyphoplasty surgery for osteoporotic vertebral fractures compared with conventional post-surgical treatment that consisted of Caltrate D (dosage of 600 mg/ day taken orally) with Rocaltrol (0.25 µg taken twice a day orally). No studies excluded individuals on medications for osteoporosis. In all other studies, the control group received no intervention.

Excluded studies

Thirty-one full journal articles, 15 abstracts, and one clinical trial registry record identified during the search were excluded in the review. Four articles (Bada 2009; Liu 2017; Sinaki 1995; Svensson 2017) were reviews or book chapters, six articles (Bautmans 2010; Borgo 2010; Hongo 2007; Jensen 2012; Lord 1996; Smith 1998) and three abstracts (Chan 2017; Kaijser Alin 2016; Marcu 2015) did not exclusively study individuals with vertebral fracture. One article did not randomize participants to groups (Sinaki 1984), one was a repeat publication of Malmros 1998, but in Danish (Malmros 1999), three were descriptive studies (Rapp 2011; Rittchen 1991; Schwinning 1992), one included an outcome that was not part of our inclusion criteria (vertebral height as a continuous measure, Webber 2003) and was a subgroup analysis from an included study (Papaioannou 2003a), and one compared standard physical therapy to casting or bracing in individuals with an acute fracture, wherein individuals less than 40 years of age were included (Stadhouders 2009). One was a systematic review of studies previously assessed for eligibility (Dusdal 2011), and one included outcomes regarding recruitment barriers in randomized controlled trials (RCTs) (Gandhi 2015). In four articles, the control group received exercises in conjunction with their

study participation (Chen 2011; Chen 2015; Karakasidou 2013; Zhong 2012). One abstract (Bergland 2012) and one full-text journal article (Olsen 2014) were identified that were secondary analyses of data from an included study (Bergland 2011). For Olsen 2014, only data on the effect of exercise on self-reported number of falls and fear of falling was extracted, which were not previously published by Bergland 2011. Three abstracts (Miyakoshi 2010; Mihailov 2011; Nemes 2014) did not include enough information to verify that it was a randomized trial or that the control group received no intervention, and no estimates of effect were provided. Three abstracts (Evstigneeva 2016; Evstigneyeva 2013; Evstigneeva 2013) were identified as secondary analyses from the study by Evstigneeva 2016. Two abstracts (Shipp 2004; Shipp 2007) were identified that included secondary analyses from the study by Gold 2004. Because the data came from a study that is included and for which we could obtain the relevant information, we report on the findings that exist in the abstract but did not include them as separate studies, clearly identifying that these data came from published abstracts only. Two articles were study protocols (Barker 2014; Giangregorio 2014c). Three conference abstracts were classified as awaiting assessment (Giangregorio 2017; Bergstrom 2016; Pompa 2016), one study was a recently in-press publication (Giangregorio 2018) and one study was an ongoing clinical trial (Stanghelle 2018).

Risk of bias in included studies

Two review authors assessed risk of bias using the scheme presented in Table 1, where YES answers referred to a low risk of bias, NO answers referred to a possibility of bias, and UNCLEAR meant it was difficult to determine whether there was potential for bias with respect to the domain in question. Discrepancies between review authors were resolved via consensus. The results are provided in the Characteristics of included studies table, Figure 2 and Figure 3.

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

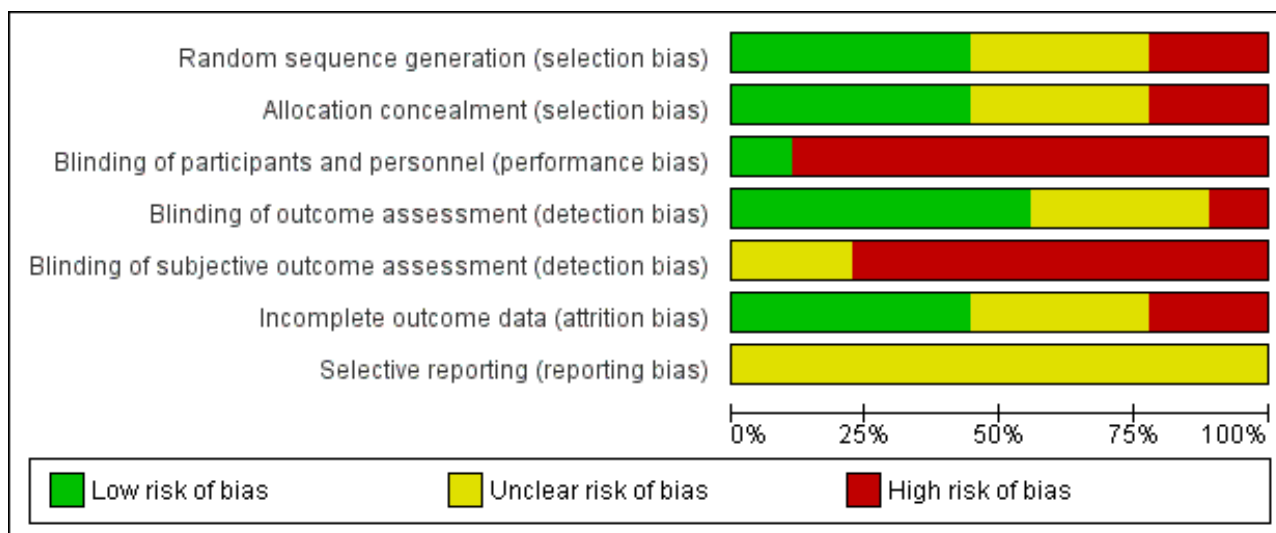


Figure 3. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Blinding of subjective outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)
Bennell 2010	+	+	-	+	-	+	?
Bergland 2011	+	+	-	+	-	+	?
Bergstrom 2011	-	-	-	?	?	-	?
Evstigneeva 2016	+	+	-	?	-	?	?
Gold 2004	+	+	+	+	?	?	?
Malmros 1998	-	-	-	+	-	+	?
Papaioannou 2003a	?	?	-	+	-	?	?
Wang 2015	?	?	-	?	-	-	?
Yang 2007	?	?	-	-	-	+	?

Allocation

In [Bennell 2010](#), [Evstigneeva 2016](#), and [Gold 2004](#), the allocation sequence was adequately generated and concealed. [Malmros 1998](#) used the drawing of sealed envelopes, so random sequence generation and allocation concealment was not ideal and it is unclear who prepared or had access to the envelopes or performed randomization. [Bergstrom 2011](#) had participants pick their allocation out of a hat; it was not clear if this was done with replacement or if the study nurse who performed randomization was the same one that performed the outcome assessment. In some cases, sequence generation or allocation concealment was not described with sufficient detail ([Papaioannou 2003a](#); [Wang 2015](#); [Yang 2007](#)).

Blinding

Blinding of outcome assessors was confirmed in all studies except for [Bergstrom 2011](#), [Evstigneeva 2016](#), and [Wang 2015](#) (unclear risk), and, and [Yang 2007](#) (high risk). Because of the nature of the intervention, neither participants nor personnel administering the intervention were blind to group allocation for any study. Additionally, blinding of subjective outcome assessment was considered at high risk of bias except for two studies that we assessed as at unclear risk ([Bergstrom 2011](#); [Gold 2004](#)).

Incomplete outcome data

In [Gold 2004](#), data were excluded from 63 of a total of 185 participants for the muscle strength outcome because of equipment problems; analyses for this outcome used a sample size of 122. [Malmros 1998](#) performed an as-treated analysis, so we chose to reanalyze the data as intention-to-treat. [Papaioannou 2003a](#) reports an intention-to-treat analysis, but it was not clear how missing data were handled. [Bergstrom 2011](#) performed an intention-to-treat analysis for the primary outcome (back extensor muscle strength) but it was not clear if an intention-to-treat or per-protocol analysis was performed for the posture outcome. [Evstigneeva 2016](#) performed an intention-to-treat analysis and used last observation carried forward method for missing data; however, two participants did not receive the allocated intervention and it was unclear whether they were included in the analysis and why they did not perform the exercises. [Wang 2015](#) initially recruited a sample of 138 participants, yet the article only examined the data of 114 patients that completed follow-up. There were no reasons provided for the lack of follow-up data in the excluded participants.

Selective reporting

Seven trials ([Bergland 2011](#); [Bergstrom 2011](#); [Evstigneeva 2016](#); [Malmros 1998](#); [Papaioannou 2003a](#); [Wang 2015](#); [Yang 2007](#)) did not report a clinical trial registration number. [Webber 2003](#) published data on vertebral height obtained in a subset of participants from the trial by [Papaioannou 2003a](#) who were not included in the main paper, however these data were evaluated for hypothesis-generating purposes and were not powered to test efficacy of the exercise intervention. [Bergstrom 2011](#) reported that there was no difference in C7 to wall distance, but the outcome is not described in the methods, and it is not clear if an intention-to-treat or per-protocol analysis was performed. [Olsen 2014](#) published a secondary analysis of new data on the effect of exercise on self-reported number of falls and fear of falling in participants, which were not reported in the original trial by [Bergland 2011](#).

[Gold 2004](#) reported data for three primary outcomes. However, the data belong to a much larger database incorporating a number of studies. Not all of the outcomes collected were reported in the paper. Among these, trunk muscle endurance (measured using the Timed Loaded Standing test) and walking endurance (measured using the Six-Minute Walk test) were reported in published abstracts ([Shipp 2004](#); [Shipp 2007](#)), but with insufficient detail to enable analysis in the current review. We have been in contact with the authors, who agreed to provide access to the data, but these were not available in time for this version of the review. The authors confirmed that the three outcomes were chosen *a priori* as the only primary outcomes. Three abstracts by [Evstigneeva 2016](#), [Evstigneeva 2013](#), and [Evstigneeva 2013](#) (see [Characteristics of excluded studies](#)) reported on a subset of outcomes (QUALEFFO-41, Timed Up and Go test, and stabilometry- Sit-to-Stand Weight Transfer, Body Weight Rising Index, and Tandem Walk and Sway) from the study by [Evstigneeva 2016](#).

The clinical trial registration number for [Bennell 2010](#) was reported, and upon reviewing the registered trial, two outcomes were listed that were not reported in the published article: standing balance on a force platform, and Human Activity Profile self-report questionnaire. Communication with the authors revealed that they were not able to process the standing balance data, so the choice to omit the data was due to logistical difficulties. The Human Activity Profile data were collected for descriptive purposes.

Other potential sources of bias

We chose to use a last observation carried forward method of imputation when reanalyzing the data from [Malmros 1998](#), which may introduce bias in the estimate of treatment effects, especially when compared to methods such as multiple imputation. However, because we were not the authors of the data, we could not verify the source of missingness to confirm whether the data were missing at random. The study reports that data were missing from five individuals (three treatment, two control) because of adverse events, and two other instances were due to participants not completing assessments, but in some cases data points were absent with no known reason.

Effects of interventions

See: [Summary of findings for the main comparison Exercise for improving outcomes after osteoporotic vertebral fracture](#)

The small number of trials and the variability in the outcomes assessed across studies, the measurement tools chosen for a given outcome, and the duration of follow-ups prevented meaningful pooling in meta-analyses for all outcomes except for the Timed Up and Go test, and the QUALEFFO-41 total score and physical function subscale score ([Summary of findings for the main comparison](#)). Studies were grouped according to primary and secondary objectives, and duration of follow-up, where a study could be represented in more than one group depending on the number of follow-up assessments: a) four to 12 weeks ([Bennell 2010](#); [Bergland 2011](#); [Malmros 1998](#); [Olsen 2014](#); [Wang 2015](#); [Yang 2007](#); [Wang 2015](#)); b) 16 to 24 weeks ([Bergstrom 2011](#); [Gold 2004](#); [Malmros 1998](#); [Papaioannou 2003a](#); [Wang 2015](#)); and c) 52 weeks ([Bergland 2011](#); [Evstigneeva 2016](#); [Olsen 2014](#); [Papaioannou 2003a](#); [Wang 2015](#)).

There were a few instances where two or more studies included comparable outcomes. However, there were a number of factors

that limited our ability to pool data. One study compared a multi-component intervention, including exercise, taping, massage and manual therapy, compared with no intervention and there was no comparison group that received the same intervention with the exception of exercise (Bennell 2010). Therefore, the effect of exercise could not be estimated and we decided not to pool these data with the other trials in the first edition of this Cochrane review (Giangregorio 2013). We could not pool data from Papaioannou 2003a because the required means or standard deviations were not provided in the report and the raw data were not available. We chose not to pool data from Gold 2004 because the required means and standard deviations were not included in the report, and because it was a cluster-trial and site was included as a covariate. Because of the limited number of studies and the variability across studies with respect to interventions, outcome measures and analysis, we have chosen to present a narrative synthesis of the findings, grouped according to outcome and follow-up time. The exceptions were the Timed Up and Go and the QUALEFFO-41 total score and physical function subscale outcome measures, where data from two trials were pooled for each. Because the intervention by Bennell 2010 combined physical therapy with exercise, it was not included in the pooled analysis in the first edition of this Cochrane Review (Giangregorio 2013). However, we conducted a sensitivity analysis with Bennell 2010 included in the meta-analysis to further assess the effects of exercise on Timed Up and Go test. Whether group mean values or mean changes are compared is indicated when known. In Bennell 2010, group means adjusted for baseline value were compared using ANCOVA for each variable. Bennell 2010; Bergland 2011; Bergstrom 2011; Papaioannou 2003a; Wang 2015, and Yang 2007 did not correct for multiple comparisons.

Major outcomes

Incident fragility fractures of the hip, vertebral or other sites

One study measured fractures as a secondary outcome (Evstigneeva 2016) (risk ratio (RR) 0.54, 95% confidence interval (CI) 0.17 to 1.71; participants = 78; very low-quality of evidence) (Analysis 1.1). Specifically, four participants in the exercise group sustained clinical vertebral and non-vertebral fractures over the 12-month study and seven participants in the control group ($P = 0.285$). It is unknown whether these fractures were symptomatic or asymptomatic.

Incident falls

One study evaluated the effects of exercise on the incidence of self-reported falls over a 12-month follow-up period (Olsen 2014) (RR 1.06, 95% CI 0.53 to 2.10; participants = 89; very low-quality of evidence) (Analysis 1.2). Specifically, Olsen 2014 reported that 27.7% of the intervention group and 26.2% of the control group had fallen over the 12-month follow-up period, which represented a reduction for both groups ($P \leq 0.001$) with no between-group differences (no data reported).

Pain

After four to 12 weeks

There were no between-group differences in the effect of exercise on pain after four weeks of exercise in Yang 2007, or after five weeks in our intention-to-treat analysis of data from Malmros 1999. After a 10-week multi-modal physical therapy intervention that included exercise, Bennell 2010 reported a between-group difference in favor of the intervention for pain on movement (mean between-group

difference adjusting for baseline value -1.8 points, 95% CI -3.5 to -0.1, $P < 0.05$) and pain at rest (mean between-group difference adjusting for baseline value -2.0 points, 95% CI -3.8 to -0.2, $P < 0.05$). Wang 2015 reported a between-group difference in favor of the intervention for pain (mean between-group difference in change score versus control -0.52, $P = 0.001$) after four weeks. There was no difference in pain between the intervention group and salmon calcitonin group (mean between-group difference in change scores -0.01 points, $P = 0.274$) after four weeks (Wang 2015).

In our intention-to-treat analysis, Malmros 1999 demonstrated a between-group difference in favor of the intervention group for pain (mean between-group change scores -1.03 points, 95% CI -1.37 to -0.69, $P = 0.013$) after a 10-week exercise program implemented by physical therapists. However, in this study, no Bonferroni correction was made to account for multiple comparisons (at least six outcomes at one or more time points). It should be noted that in the original paper, Malmros 1999 did not report whether the between-group differences were observed for pain at any time point, but stated across all time points (five weeks, 10 weeks, 22 weeks) that a quote: "...difference was found between the course of values from the two study groups ($P = 0.02$)."

After 16 to 24 weeks

For pain with activities, no between-group difference was observed in Gold 2004 (mean between-group difference in change scores -0.03, $P = 0.640$). The difference in change scores that was reported from our intention-to-treat analysis of data from Malmros 1999 after 10 weeks of exercises was no longer observed (mean between-group difference in change scores -0.73 points, 95% CI -1.0 to 0.36, $P = 0.09$) at 22 weeks follow-up when contact with the physical therapist had ceased. Wang 2015 reported a between-group difference in favor of the intervention for pain (mean between-group difference in change score versus control -0.72 points, $P = 0.001$ and versus salmon calcitonin treatment -0.45 points, $P = 0.001$) after 24 weeks.

After 52 weeks

Wang 2015 reported a between-group difference in favor of the intervention for pain (mean between-group difference in change score versus control -1.28 points, $P = 0.001$ and versus salmon calcitonin treatment -0.97 points, $P = 0.001$) after 52 weeks.

Physical performance: performance-based measures

After four to 12 weeks

Maximum walking speed over 20 meters in Bergland 2011 improved after three months of exercise (mean change score -1.3 seconds, 95% CI -2.0 to -0.6 for the intervention group versus 0.6 seconds, 95% CI -0.3 to 1.4 for the control group, effect size 0.5, $P < 0.001$) (Analysis 2.1). Functional reach also improved (mean change score 1.7 cm, 95% CI 0.1 to 3.1 for the intervention group versus -2.2 cm, 95% CI -3.8 to -0.7 for the control group, effect size = 0.6, $P < 0.001$). Performance on the Timed Up and Go test improved in the intervention group compared to the control group in two studies after exercising for four weeks (mean change score 11.87 seconds, SD = 2.13 for the intervention group versus 14.4 seconds, SD = 3.08 for the control group, $P < 0.05$, Yang 2007) and after exercising for 12 weeks (mean change score -0.5 seconds, 95% CI -0.9 to 0.1 for the intervention group versus 0.4 seconds, 95% CI -0.2 to 1.1 for the control group, effect size 0.2, $P < 0.026$, Bergland 2011). In contrast, one study showed no between-group difference in performance

on the Timed Up and Go test after 10 weeks of follow-up (mean between-group difference adjusting for baseline value 0.5 seconds, 95% CI -1.6 to 0.6, [Bennell 2010](#)). When data from two of the studies ([Bergland 2011](#); [Yang 2007](#)) were pooled, there was a small effect of exercise on performance on the Timed Up and Go test (MD -1.13 seconds, 95% CI -1.85 to -0.42, participants = 119) ([Analysis 2.2](#)). When the data from the two studies in the meta-analysis were pooled with [Bennell 2010](#) for a sensitivity analysis, the effect of exercise on Timed Up and Go test performance did not change (MD -1.09 seconds, 95% CI -1.78 to -0.40; participants = 139; moderate-quality evidence) ([Table 3](#)). Time to get up from a supine position improved relative to controls after four weeks of exercise (mean change score 4.87 seconds, SD = 1.17 for the intervention group versus 7.73 seconds, SD = 1.65 for the control group, $P < 0.001$, [Yang 2007](#)).

After 16 to 24 weeks

No between-group difference was observed in [Papaioannou 2003a](#) for Timed Up and Go test performance (mean between-group difference in change scores -0.01, 95% CI -1.58 to 1.56, $P = 0.99$).

After 52 weeks

Maximum walking speed over 20 meters in [Bergland 2011](#) still improved at 12-month follow-up of a three-month exercise program (mean change score -0.9 seconds, 95% CI -1.4 to 0.3 for the intervention group versus 0.6 seconds, 95% CI -0.6 to 1.8 for the control group, effect size = 0.4, $P = 0.019$). [Evstigneeva 2016](#) reported an improvement in Timed Up and Go test performance after 52 weeks (mean change score -0.65 seconds, 95% CI -1.31 to -0.10 for intervention group and mean change score 0.29 seconds, 95% CI -0.37 to 0.95 for control group, $P = 0.020$). No between-group difference was observed in [Papaioannou 2003a](#) ($P > 0.10$; point estimates and 95% CI not reported), whereas [Bergland 2011](#) observed a between-group difference for Timed Up and Go test performance after 52 weeks (mean change score -0.6 seconds, 95% CI -1.0, -0.2 for the intervention group versus 0.2 seconds, 95% CI -0.4, 0.7 for the control group, effect size = 0.3, $P = 0.021$). Functional reach was no longer different between groups at 12-month follow-up in [Bergland 2011](#) (mean change score 1.1 cm, 95% CI -0.7 to 2.7 for the intervention group versus -0.3 cm, 95% CI -1.0 to -1.6 for the control group).

Physical performance: self-reported physical function questionnaires

After four to 12 weeks

[Bennell 2010](#) reported no between-group difference in activity restriction (mean between-group difference adjusting for baseline value -1.8 points, 95% CI -3.9 to 0.3), but did report an improvement in the physical function subscale of the QUALEFFO-41 (mean between-group difference adjusting for baseline value -4.8 points, 95% CI -9.2 to -0.5, $P < 0.05$) on a zero to 100 scale, where lower scores indicate better physical function. [Bergland 2011](#) found no improvement in the physical function subscale of the QUALEFFO-41 (mean change score -2.1 points, 95% CI -4.9 to 0.8 for the intervention group versus -0.6 points, 95% CI -2.6 to 1.5 for the control group, $P = 0.40$). When the data from [Bergland 2011](#) and [Bennell 2010](#) were pooled, there was a small effect of exercise on the physical function subscale of the QUALEFFO-41 (MD -2.84 points, 95% CI -5.57 to -0.11; very low-quality of evidence) ([Analysis 2.3](#)). No effect of exercise was observed on self-reported physical function using the Oswestry questionnaire in our intention-to-

treat analysis after 10 weeks by [Malmros 1999](#), which is similar to what was reported in the original paper. [Wang 2015](#) reported no between-group difference in self-reported physical function using the Oswestry Disability Index questionnaire after four weeks of the intervention (mean between-group difference in change scores versus control -1.9 points, $P = 0.519$ and versus salmon calcitonin treatment -0.18 points, $P = 0.445$).

After 16 to 24 weeks

No between-group difference was observed in [Papaioannou 2003a](#) for physical function (mean between-group difference in change scores 0.22 points, 95% CI -0.08 to 0.52, $P = 0.15$) or activities of daily living subscales of the OQLQ (mean between-group difference in change scores 0.17 points, 95% CI -0.09 to 0.43, $P = 0.16$). [Wang 2015](#) reported a between-group difference in favor of the intervention for self-reported physical function using the Oswestry Disability Index questionnaire (mean between-group difference in change score versus control -7.39 points, $P = 0.001$ and versus salmon calcitonin treatment -5.35 points, $P = 0.001$) after 24 weeks.

After 52 weeks

Although [Bergland 2011](#) reported no between-group difference in the change score for the physical function subscale of the QUALEFFO-41 after the 12-week intervention, it was different between groups at one-year follow-up (mean change score -2.5 points, 95% CI -5.0 to -0.03 for the intervention group versus -1.0 points, 95% CI -1.5 to 3.4 for the control group, effect size 0.3, $P = 0.047$). A between-group difference in the activities of daily living subscale of the OQLQ was observed after one year of exercise (mean between-group difference in change scores 0.34 points, 95% CI -0.11 to 0.79, $P = 0.04$) in [Papaioannou 2003a](#), but not in the physical function subscale (mean between-group difference in change scores 0.16 points, 95% CI -0.35 to 0.68, $P = 0.18$). [Wang 2015](#) reported a between-group difference in favor of the intervention for self-reported physical function using the Oswestry questionnaire (mean between-group difference in change score versus control -14.2 points, $P = 0.001$ and versus salmon calcitonin treatment -12.6 points, $P = 0.001$) after 52 weeks. [Evstigneeva 2016](#) reported between-group differences in self-reported physical function subscales from the QUALEFFO-41 questionnaire (mean change score for jobs around the house subscale -9.8 points, 95% CI -13.6 to -5.9 for the intervention group versus -0.3, 95% CI -4.1 to 3.6 for the control group, $P = 0.002$ and mean change score for mobility subscale -7.3 points, 95% CI -10.6 to -3.9 for the intervention group versus 6.2 points, 95% CI 2.1 to 10.3 for the control group, $P < 0.001$). However, there was no between-group difference in the self-reported physical function-activities of daily living subscale from the QUALEFFO-41 questionnaire (mean change score -4.7 points, 95% CI -8.8 to -0.6 versus 0.5 points, 95% CI -4.1 to 5.1, $P = 0.076$) ([Evstigneeva 2016](#)).

Health-related quality of life: disease-specific scales

After four to 12 weeks

[Bennell 2010](#) reported no differences in QUALEFFO-41 total score (mean between-group difference in change scores -7.1 points, 95% CI -14.9 to 0.8) or in any subscales other than physical function (see Physical Function - Self-report Questionnaires section above), after 10 weeks of a multi-modal physical therapy intervention that included exercise. [Bergland 2011](#) reported no improvement in the total score (mean change score -2.1 points, 95% CI -4.2 to -0.02 for the intervention group versus 0.2 points, 95% CI -2.2 to 2.5 for the

control group) or any of the subscales of the QUALEFFO-41, with the exception of the mental function subscale (mean change score -5.1 points, 95% CI -8.1 to -1.6 for the intervention group versus 2.6 points, 95% CI -1.8 to 6.9 for the control group, effect size 0.6, $P < 0.006$). When the data from [Bergland 2011](#) and [Bennell 2010](#) were pooled, there was a small effect of exercise on the QUALEFFO-41 total score (MD -3.24 points, 95% CI -6.05 to -0.43; very low-quality of evidence) ([Analysis 2.4](#)).

After 16 to 24 weeks

Between-group differences favoring exercise were observed in several OQLQ subscales in [Papaioannou 2003a](#): symptom (mean between-group difference in change scores 0.44 points, 95% CI 0.16 to 0.73, $P < 0.003$); emotion (mean between-group difference in change scores 0.34 points, 95% CI 0.02 to 0.66, $P < 0.01$); and leisure/social (mean between-group difference in change scores 0.39 points, 95% CI -0.02 to 0.81, $P = 0.03$).

After 52 weeks

Although no improvement in QUALEFFO-41 total score was observed after a three-month exercise program, [Bergland 2011](#) reported a between-group difference at 12-month follow-up (mean change score -3.3 points, 95% CI -5.2 to -1.3 for the intervention group versus -0.4 points, 95% CI -2.0 to -2.7 for the control group, effect size 0.3, $P < 0.019$). The between-group difference in the mental function subscale of the QUALEFFO-41 observed at three months was maintained at 12-month follow-up (mean change score -2.6 points, 95% CI -6.2 to -0.9 for the intervention group versus 2.7 points, 95% CI -1.0 to 6.5 for the control group, effect size 0.4, $P < 0.04$), and between-group difference in the pain subscale (mean change score -13.6 points, 95% CI -19.3 to -7.8 for the intervention group versus -1.8 points, 95% CI -7.7 to 4.1 for the control group, effect size 0.5, $P < 0.005$) was also observed at this time point. A between-group difference favoring exercise was observed in [Papaioannou 2003a](#) in the OQLQ symptom subscale (total OQLQ data not reported): symptom (mean change score 0.38 points, 95% CI -0.05 to 0.81, $P < 0.02$). [Papaioannou 2003a](#) reported that individuals who completed at least three days of exercise a week had greater improvements in OQLQ symptoms score ($P = 0.017$) and total OQLQ score ($P = 0.048$). [Evstigneeva 2016](#) found improvements in the QUALEFFO-41 total score (mean change score -5.8 points, 95% CI -7.8 to -3.8 for the intervention group versus 3.1 points, 95% CI 1.3 to 4.9 for the control group, $P < 0.001$), pain (mean change score -13.5 points, 95% CI -18.5 to -8.5 for the intervention group versus -0.4 points, 95% CI -5.7 to 4.9 for the control group, $P = 0.001$), social function (mean change score -3.7 points, 95% CI -8.5 to 1.0 for the intervention group versus 3.5 points, 95% CI -0.7 to 7.8 for the control group, $P = 0.012$), and general health perception (mean change score -6.0 points, 95% CI -10.8 to -1.3 for the intervention group versus 8.1 points, 95% CI 4.0 to 12.2 for the control group, $P < 0.001$).

Health-related quality of life: generic scales

After four to 12 weeks

[Bennell 2010](#) reported no between-group difference in Assessment of Quality of Life (AQoL) scores (mean between-group difference adjusting for baseline value 0.10 points, 95% CI -0.04 to 0.24). [Bergland 2011](#) reported a between-group difference for General Health Questionnaire (GHQ) total score (mean change score -3.7 points, 95% CI -5.5 to -1.9 in the intervention group versus -0.2

points, 95% CI -2.1 to 1.7 in the control group, effect size 0.4, $P < 0.009$).

After 16 to 24 weeks

No between-group differences were observed in [Papaioannou 2003a](#) for the total score (mean between-group difference in change scores 0.55, 95% CI -1.81 to 2.91, $P = 0.54$) or subscales of the Sickness Impact Profile - physical (mean between-group difference in change scores 0.80, 95% CI -1.52 to 3.13, $P = 0.94$) and psychological/social (mean between-group difference in change scores 0.09, 95% CI -3.21 to 3.41, $P = 0.07$).

After 52 weeks

Although total GHQ scores were improved relative to controls after three months of exercise in [Bergland 2011](#), the between-group difference was no longer observed at 12-month follow-up (mean change score -2.8 points, 95% CI -4.6 to -1.0 for the intervention group versus -1.1 points, 95% CI -2.8 to 0.7 for the control group).

Adverse events

Adverse events were reported in the results section of five trials ([Bennell 2010](#); [Bergstrom 2011](#); [Evstigneeva 2016](#); [Gold 2004](#); [Papaioannou 2003a](#)), and one trial ([Malmros 1999](#)) indicated that there were five adverse events unrelated to study participation, but did not describe them or their severity. [Table 2](#) summarizes the reported adverse events. Four events, including two fractures, were directly attributable to exercise (very low-quality of evidence).

Minor outcomes

Posture

After four to 12 weeks

[Bennell 2010](#) reported no between-group difference in degrees of thoracic kyphosis (mean between-group difference adjusting for baseline value -2.9 degrees, 95% CI -7.9 to 2.1).

After 16 to 24 weeks

[Bergstrom 2011](#) reported no between-group difference in thoracic kyphosis after the intervention ($P = 0.90$; point estimates and 95% CI not reported).

After 52 weeks

[Evstigneeva 2016](#) reported no between-group difference in occiput-to-wall distance (mean change score -0.71 cm, 95% CI -1.31 to -0.10 in the intervention group versus -0.09 cm, 95% CI -0.77 to 0.59 in the control group, $P = 0.208$).

Muscle function

After four to 12 weeks

An effect of a multi-modal physical therapy intervention including exercise was observed for trunk muscle endurance in [Bennell 2010](#), measured via Timed Loaded Standing (mean between-group difference adjusting for baseline value 46.7 seconds, 95% CI 16.1 to 77.3, $P < 0.05$). No effect of exercise on back extensor muscle strength was observed after five or 10 weeks by [Malmros 1999](#); the P value reported in the original paper for the 10 week comparison was $P = 0.09$.

After 16 to 24 weeks

For trunk extension muscle strength, a between-group difference in favor of exercise was observed in [Gold 2004](#) (mean between-group difference in change scores 10.68 pounds, 95% CI 6.98 to 14.39, $P < 0.001$, $n = 122$, subgroup of total sample $n = 185$). From an abstract ([Shipp 2007](#)) using the study sample from [Gold 2004](#), there was no between-group difference in change scores for trunk and arm muscle endurance (point estimates and 95% CI not reported). [Bergstrom 2011](#) reported no difference in back extensor muscle strength when control and intervention groups were compared in intention-to-treat analyses (mean change score 254 ± 85 N in the control group versus 302 ± 108 N in the intervention group, $P = 0.74$). They reported a between-group difference in back extensor muscle strength ($P = 0.029$) in a per-protocol analysis, where eight individuals who dropped out or were not compliant with exercise or control activities were excluded, and adjustment was made for baseline differences.

Balance

After 12 weeks or less

No studies measured balance at this time point.

After 16 to 24 weeks

In [Papaioannou 2003a](#), tests of postural sway revealed a between-group difference in favor of exercise for the range of displacement during the eyes closed condition (mean between-group difference in change scores -0.80 cm, 95% CI -1.45 to -0.15 , $P = 0.01$), but not for any of the other measured postural sway displacement of velocity variables in the eyes closed condition, or for any postural sway variables in the eyes open condition. No effect of exercise on postural sway was observed after 22 weeks in our analysis of data from [Malmros 1999](#); in the original paper, they reported a trend towards an improvement in balance with exercise.

After 52 weeks

[Papaioannou 2003a](#) reported that displacement in lateral and anteroposterior directions and velocity of movement (measured with a force plate) improved in the intervention group relative to the control group ($P < 0.01$), but no data were provided, and it is not clear if this is for the eyes open or eyes closed condition. [Evstigneeva 2016](#) demonstrated improvements in sit-to-stand weight transfer (mean change score -0.24 seconds, 95% CI -1.12 to 0.64 in the intervention group versus 0.43 seconds, 95% CI 0.13 to 0.73 in the control group, $P = 0.010$) and tandem walk and sway tests (mean change score -0.08 degrees/second, 95% CI -1.62 to 1.47 in the intervention group versus 1.72 degrees/second, 95% CI 0.15 to 3.30 in the control group, $P = 0.029$). However, no between-group differences were observed in [Evstigneeva 2016](#) for sit-to-stand left/right weight symmetry (mean change score in the intervention group 0.60% , 95% CI -1.52 to 2.72 versus 1.97% , 95% CI -0.76 to 4.69 in the control group, $P = 0.142$) and weight-bearing squat tests (mean change score of 1.40% , 95% CI -1.06 to 3.86 in the intervention group versus 0.53% , 95% CI -1.98 to 3.03 in the control group, $P = 0.638$).

Bone mineral density of the lumbar spine or hip measured using DXA

After 52 weeks

[Papaioannou 2003a](#) reported no effect of thrice-weekly home exercise for one year on lumbar spine or femoral neck bone mineral density (no data reported). [Wang 2015](#) reported a between-group difference in favor of the intervention for lumbar bone mineral density (mean between-group difference in change score versus control 0.038 g/cm², $P = 0.005$ and versus salmon calcitonin treatment 0.037 g/cm², $P = 0.042$) after 52 weeks.

Fear of falling

After four to 12 weeks

[Olsen 2014](#) reported a between-group difference in favor of the intervention group for Falls Self-Efficacy Scale-I score (mean between-group difference in change scores 3.3 , 95% CI 1.1 to 5.4 , effect size 0.4 , $P = 0.004$) after three months.

After 52 weeks

[Olsen 2014](#) reported a between-group difference in favor of the intervention group for Falls Self-Efficacy Scale-I score (mean between-group difference in change scores 5.4 , 95% CI 3.0 to 7.9 , effect size 0.7 , $P < 0.001$) after 12 months.

Patient assessment of global success

None of the included studies measured patient assessment of global success at any of the time points.

DISCUSSION

Summary of main results

We have summarized the findings from nine published randomized controlled trials, including two new trials in this update of our 2013 Cochrane Review ([Giangregorio 2013](#)). The limited number of studies and the diversity in outcomes reported, measurement tools used for a given construct, and duration of follow-up prevented any meaningful pooling of data for most outcomes. Among studies that had similar outcomes and follow-up durations, there were contradictory findings.

No trials measured incident fractures, falls, or adverse events as primary outcomes. Only one study each evaluated incident fractures ([Evstigneeva 2016](#)) and self-reported incident falls ([Olsen 2014](#)) as a secondary outcome with no evidence of a difference in effect between intervention and control groups and the 95% confidence intervals included huge benefits and huge harms. Some adverse events occurred that were directly associated with the interventions, including four fractures. Most studies did not report if and how adverse event data were collected in the methodology section, and adverse events were reported inconsistently. It would be important for future studies to systematically collect and report adverse events as a study outcome, and identify those related to the intervention and their severity.

Individual studies demonstrated between-group differences in the effect of exercise on major outcomes including pain, maximum walking speed, Timed Up and Go test, and disease-specific quality of life (either as a total score or as one of the subscales of a quality of life tool); yet the differences were not clinically important.

There were also individual studies reporting no between-group differences in favor of exercise on some of these outcome measures representing the same construct; this was true for pain, Timed Up and Go test, and disease-specific quality of life. The effects of exercise on performance-based measures are promising, but the magnitude of the effects was small (between-group differences ~1 second for Timed Up and Go and ~2 seconds for maximum walking speed over 20 meters) and not clinically meaningful (clinically important improvements in Timed Up and Go typically range from 1.4 to 3.4 seconds in other populations with chronic musculoskeletal conditions) (Gautschi 2017; Wright 2011). Pooled analyses of two trials revealed a between-group difference in favor of exercise for Timed Up and Go test (Bergland 2011; Yang 2007). These findings should be viewed with caution given two other trials that could not be included in the original pooled analysis did not find between-group differences. However, a sensitivity analysis including data from one of the trials excluded from the meta-analysis in the first edition of this review (Giangregorio 2013), a multi-component intervention, including exercise, taping, massage and manual therapy, compared with no intervention (Bennell 2010), did not modify the observed effects of exercise on Timed Up and Go test performance. Individual studies that demonstrated clinically important differences in favor of exercise for pain (>1-point improvement or 15% change) (Salaffi 2004), also defined the presence of pain as an inclusion criterion (Bennell 2010; Malmros 1998; Wang 2015).

Only four studies implemented an intervention or follow-up of 52 weeks (Bergland 2011; Evstigneeva 2016; Papaioannou 2003a; Wang 2015), and small between-group differences in favor of exercise were reported for some major outcome measures in these studies, including maximum walking speed, Timed Up and Go test, pain, and disease-specific health-related quality of life (total score or select subscales). However, there were also reports from certain studies showing a lack of clinically important effects of exercise on Timed Up and Go test, self-reported physical function from the Oswestry Disability Index, and disease-specific quality of life subscales. Findings for minor outcomes (time to get up from a supine position, trunk muscle endurance, back extensor muscle strength, range of displacement during a postural sway test with eyes closed or in tandem stand position, functional mobility during a sit-to-stand weight transfer, generic quality of life, lumbar spine bone mineral density, and fear of falling) are less certain, due to the relatively low number of studies and participants.

Adherence to exercise varied across studies; adherence appeared to be higher among studies that included supervised, patient-specific assessment and prescription, even if it was intermittent. In the two studies with long-term follow-up, adherence decreased when the level of supervision decreased. Despite identifying several new full-text journal publications and abstracts since the publication of our original review (Excluded studies), only two new trials met all of our inclusion criteria (Evstigneeva 2016; Wang 2015), and we extracted new data on falls and fear of falling from a secondary analysis (Olsen 2014), which were not previously reported in the original trial by Bergland 2011. However, we did identify three conference abstracts (Bergstrom 2016; Giangregorio 2017; Pompa 2016), one publication in press (Giangregorio 2018), and one ongoing trial (Stanghelle 2018) to be evaluated for inclusion in future editions of this review.

Overall completeness and applicability of evidence

The quality of reporting has improved since the advent of the CONSORT guidelines, but there is room for further improvement. The group means and standard deviations, point estimates of effect, 95% confidence intervals and P values were not consistently reported for all outcomes, which would be helpful if future versions of this review are able to pool data. In some cases between-group differences are reported as actual differences and in other cases they are reported as change scores. Ideally one is chosen as a primary outcome but both are reported, or the raw data are archived so they can be shared with researchers performing knowledge syntheses. All of the exercise programs included muscle strengthening; yet, the specific exercise regimens were not always adequately described (e.g. which muscles were targeted or magnitude, progression of resistance or number of repetitions). Regarding reporting of the exercise intervention, Bennell 2010 included a table listing details of individual exercises, which is ideal for readers who will attempt to apply the research to practice; for future studies, this could be included in the body of the paper or as an appendix. It should be mandatory to report the protocol for progressing the exercises. There have been only two studies to date that included men with vertebral fracture. More studies examining the effect of exercise in men with vertebral fracture as a single group or as a subgroup of a larger trial including both genders are necessary to make any conclusive inferences.

Quality of the evidence

The small number of studies, variability in study design, and high or unclear risk of bias due to some methodological components limits the conclusions that can be made to date regarding the effectiveness of exercise for women with vertebral fracture. All studies were rated as high risk of bias for "blinding of participants and personnel" because the nature of exercise limits the ability to blind participants or those implementing the intervention. All studies were rated as unclear or high risk for at least one other criterion in the assessment of risk of bias (Characteristics of included studies and Figure 3), and the quality of studies was quite variable. In many cases the protocol was not publicly available, as this did not become a requirement until more recently. The imprecision in the estimates of effect for many outcomes (because of inconsistent findings across studies) also contributes to our conclusion that the body of evidence around the benefits of exercise for most outcomes after vertebral fracture is of very low quality.

Potential biases in the review process

We restricted the review to randomized controlled trials. An inclusion criterion for all trials was a history of vertebral fracture, but the way vertebral fracture was defined varied between studies (i.e. clinical versus morphometric, variable definitions of morphometric fracture), which may create dissimilar study populations and variable effects, and may affect generalizability. For example, Malmros 1998 and Bennell 2010 included only women with a painful vertebral fracture at baseline, whereas Gold 2004 defined fractures based on morphometry; those with symptomatic fractures may have a different response to exercise for outcomes such as pain or quality of life. Our preference is to report symptomatic and asymptomatic fractures separately, since asymptomatic radiographic fractures may not be clinically relevant vertebral fractures. However, Evstigneeva 2016 did not describe

how incident fragility fractures were confirmed in their study or whether they included symptomatic or asymptomatic fractures, or both. When we were able to communicate with authors or review a clinical trial registration, we uncovered instances of incomplete outcome reporting (Bennell 2010; Gold 2004). The reasons provided by the authors for the incomplete outcome reporting (i.e. difficulty with equipment, reporting the *a priori* defined outcomes only) did not appear to be attempts to hide information. However, it supports the need for clinical trial registration, and for thoroughness in evaluations of risk of bias related to incomplete outcome reporting. There were a few instances of incomplete outcome data (i.e. as-treated analysis, subgroup analysis on participants for whom outcomes could be assessed, no information about how missing data were treated).

It is very possible that the lack of effect or contradictory findings observed were related to studies being underpowered, as well as variability in participants, exercise interventions, co-interventions, comparators, follow-up times and outcomes chosen. For example, the multi-modal physical therapy intervention used by Bennell 2010 including manual therapy, massage, taping and exercises may have had interactive effects, and it may not be appropriate to pool this trial with others that intervened with exercise only. Four trials were excluded because of the heterogeneity of study participants (Bautmans 2010; Hongo 2007; Jensen 2012; Smith 1998). The participants in Bautmans 2010 and Hongo 2007 were women with postmenopausal osteoporosis, and individuals with vertebral fractures at baseline were not excluded (unless they were recent or symptomatic), so it is possible that a subgroup of participants in each of the studies were representative of the participants who were of interest in the current review. We chose not to include these studies to avoid heterogeneity of effects, as those without vertebral fracture may be different with respect to response to exercise than those with vertebral fracture. Bautmans 2010 reported an improvement in thoracic kyphosis (measured with Spinal-Mouse) in the intervention group compared with controls, but no between-group difference in disease-specific quality of life (QUALEFFO-41) or back pain (Visual Analog Scale); the intervention included postural taping, spinal mobilization, and postural exercises over 18 sessions with a physical therapist. Hongo 2007 included women with and without vertebral fractures, and has the same authors as those in Miyakoshi 2010. However, although the outcomes and authors are similar, the interventions described in each trial are slightly different with respect to exercises chosen, follow-up time, and exercise frequency, so it is likely that they are independent trials. No between-group difference in spinal range of motion was reported in Hongo 2007; the exercise intervention was one set of 10 repetitions of back extension exercises performed in prone lying five days a week for four months. Hongo 2007 and Miyakoshi 2010 reported a between-group difference in favor of exercise for back extensor muscle strength and quality of life (Japanese Osteoporosis Quality of Life questionnaire).

Agreements and disagreements with other studies or reviews

The observed benefit of exercise for walking speed is consistent with the findings of a recent meta-analysis demonstrating that exercise interventions in frail older adults can improve walking speed (Chou 2012). A systematic review entitled "*Effects of therapeutic exercise for persons with osteoporotic vertebral fractures: a systematic review*" was published in 2010 (Dusdal 2010).

We are in agreement with the authors' position that no definitive conclusions can be drawn about the effects of exercise in this population, except for its effect on physical performance, and that future research is needed. Several differences between the reviews should be acknowledged. We limited our review to randomized controlled trials of exercise where all of the participants had to have at least one osteoporotic vertebral fracture. Dusdal 2010 included non-randomized trials, and specified that to be included 50% of the study participants had to have had at least one vertebral fracture. Their criteria resulted in the inclusion of an additional three studies (Bautmans 2010; Hongo 2007; Sinaki 1984) and one abstract (Smith 1998) that we had officially excluded (see Excluded studies). We excluded an additional article that may have met their eligibility criteria (Webber 2003). We included four full-text journal publications that they did not; Bergland 2011 was published around the same time as their review, and Bergstrom 2011, Evstigneeva 2016, and Wang 2015 were published after it. We identified information from secondary data analyses (Shipp 2004; Shipp 2007) under the trial by Gold 2004, albeit we did exclude the abstracts as stand-alone contributions to the review. We chose to re-analyze the data from Malmros 1998 because an intention-to-treat analysis was not published and our re-analysis resulted in a different interpretation of the findings of that study and provided data regarding the effects of the intervention at five, 10 and 22 weeks of follow-up. We did not examine the effects of exercise on range of motion, but it was reported in Dusdal 2010. We also included new data on self-reported falls and fear of falling from Olsen 2014, which is a secondary analysis of the Bergland 2011 study. Finally, we emphasized between-group comparisons only, whereas Dusdal 2010 discussed both within- and between-group comparisons. This is relevant because in some cases there were within-group improvements but no between-group differences, and this may affect the interpretation of the efficacy of exercise.

AUTHORS' CONCLUSIONS

Implications for practice

There was moderate-quality evidence that exercise probably improves physical performance in individuals with vertebral fracture. No other definitive conclusions can be made regarding the benefits of exercise on incident fragility fractures, falls, adverse events, and other patient-reported outcome measures (pain, quality of life) for individuals with a history of vertebral fracture. Individual trials did report benefits for some minor outcome measures, including balance, back extensor muscle strength, trunk muscle endurance, bone mineral density, and fear of falling. These findings should be interpreted with caution, especially given the heterogeneity in the direction and size of effects. The effects on physical performance (i.e. improvements in maximum walking speed of 1.3 seconds over 20 meters, or a one second improvement in Timed Up and Go test) were small. Despite the between-group differences for disease-specific quality of life and self-reported physical function (improvements of 3 points out of 100 points), our findings likely do not indicate any clinically meaningful improvements in these outcomes after exercise intervention in individuals with vertebral fractures. Because all of the exercise programs included muscle strengthening, it is reasonable to suggest that if an individual would like to achieve the potential benefits suggested by any of the trials, muscle strengthening must be included in the exercise prescription. Back extensor muscle strength was often emphasized, but it cannot be confirmed if

potential benefits are attributable to the inclusion of progressive resisted back extensor muscle exercises. It does point to some consensus among researchers in the field that this is an important impairment to target. Supervision may affect adherence. Many of the exercise programs were center-based and supervised by a physical therapist, so it is unknown whether an intervention would be more or less effective in a different setting.

Implications for research

The experience of vertebral fracture confers a unique effect on quality of life, fall risk, posture, and pain (Adachi 2001; Cauley 2000; Kado 2007). It may not be appropriate to generalize results of exercise studies conducted in older adults without vertebral fractures to those with vertebral fractures because the safety and efficacy of exercise may differ, and it may be necessary to tailor the exercise to address impairments unique to this population. There is a need for more adequately powered, high-quality randomized controlled trials to inform exercise prescription in individuals at high risk of fracture. All but two trials (Gold 2004; Wang 2015) had less than 100 participants, and only five trials had more than 60 participants. Future research should account for adherence and attrition in sample size calculations: the adherence to trials of thrice-weekly exercise for six months or greater is often less than 60%, and the dropout rate can be 20% or greater, suggesting that there is a need to develop strategies for enhancing adherence or facilitating behavior change. Only two studies (Bergland 2011; Evstigneeva 2016) reported a sample size calculation. Among women with vertebral fractures on alendronate therapy, 8.0% and 11.9% experienced a new vertebral or a clinical non-vertebral fracture, respectively, over 2.9 years of follow-up, versus 15% and 14.7% respectively in the placebo group (Black 1996). If all participants were on anti-osteoporosis therapy, we would estimate that 6% of participants would experience a new vertebral fracture or non-vertebral fracture in one year (Black 1996), and to observe a fracture risk reduction of 30%, the larger trial would require ~4400 participants. However, in untreated women we anticipate 10% would experience any fracture, resulting in a required sample size of 2500 individuals. Some patients choose not to accept therapy, or are nonadherent, so the required sample size may lie between these two estimates. Outcomes important to people with osteoporotic vertebral fracture and the health systems, such as falls, fractures, disability and health services costs,

should be considered for inclusion in future trials. Adverse events should be assessed in both groups (intervention and comparator) to ensure that the risks do not outweigh any benefits of an exercise intervention. Future work should also consider whether the number, location or severity of vertebral fractures, vitamin D levels, pain or bisphosphonate use at baseline modify the effect. From a theoretical perspective, many of the interventions target back extensor muscle endurance or back extensor muscle strength to reduce hyperkyphosis and resist forward flexion of the spine. Future research should specifically state whether the intervention targets muscle strength, endurance or both, and assess change in both strength and endurance outcome measures. Furthermore, it may be prudent to investigate whether change in strength or endurance account for change in other outcome measures such as posture and pain, and quantify these effects on vertebral fractures.

ACKNOWLEDGEMENTS

The project was initially supported in part by the Canadian Institutes for Health Research (CIHR) grant for the senior author under the CIHR Randomized Controlled Trials Mentoring Program. We would like to thank Elizabeth Ghogomu, Jordi Pardo, Lara Maxwell and the Cochrane Musculoskeletal Group for their support with the preparation of this manuscript. Many thanks to Louise Falzon, Tamara Rader, and Jackie Stapleton for contributing their expertise to the design and conduct of the literature search. We are appreciative of Carly Skidmore for her assistance with the literature review and data extraction, and with preparing figures for the original published review. Thank you to Drs. Malmros, Bergland, Bennell, Summers, Mortensen, and Lesnyak for their assistance with our queries. Thank you to Ann Cranney and Richard Cook for their contributions to the protocol.

L. Giangregorio held a CIHR New Investigator Award and an Early Researcher Award from the Ontario Ministry of Research and Innovation which contributed to the available time and personnel to complete the review. N. MacIntyre and L. Giangregorio are past recipients of CIHR Randomized Controlled Trials Mentoring Program awards. L. Thabane and A. Papaioannou are CIHR clinical trials mentors from the CIHR Randomized Controlled Trials Mentoring Program. A. Papaioannou is the CIHR-Eli Lilly Chair in Osteoporosis and Fracture Prevention. J. Gibbs held a CIHR Fellowship Award.

REFERENCES

References to studies included in this review

Bennell 2010 {published data only}

Bennell KL, Matthews B, Greig A, Briggs A, Kelly A, Sherburn M, et al. Effects of an exercise and manual therapy program on physical impairments, function and quality-of-life in people with osteoporotic vertebral fracture: a randomised, single-blind controlled pilot trial. *BMC Musculoskeletal Disorders* 2010;**11**:36.

Bergland 2011 {published data only}

* Bergland A, Thorsen H, Karesen R. Effect of exercise on mobility, balance, and health-related quality of life in osteoporotic women with a history of vertebral fracture: a randomized, controlled trial. *Osteoporosis International* 2011;**22**(6):1863-71.

Olsen CF, Bergland A. The effect of exercise and education on fear of falling in elderly women with osteoporosis and a history of vertebral fracture: results of a randomized controlled trial. *Osteoporosis International* 2014 Aug;**25**(8):2017-25.

Bergstrom 2011 {published data only}

Bergstrom I, Bergstrom K, Kronhed A, Karlsson S, Brinck J. Back extensor training increases muscle strength in postmenopausal women with osteoporosis, kyphosis and vertebral fractures. *Advances in Physiotherapy* 2011;**13**(3):110-7.

Evstigneeva 2016 {published data only}

* Evstigneeva L, Lesnyak O, Bultink IEM, Lems WF, Kozhemyakina E, Negodaeva E, Guselnikova G, Belkin A. Effect of twelve-month physical exercise program on patients with osteoporotic vertebral fractures: a randomized, controlled trial. *Osteoporosis International* Aug 2016;**27**(8):2515-24. [DOI: [10.1007/s00198-016-3560-4](https://doi.org/10.1007/s00198-016-3560-4)]

Evstigneeva L, Lesnyak O, Negodaeva E, Guselnikova G, Kozhemyakina E. A randomized, prospective study of the effects of exercise on quality of life in postmenopausal women with vertebral fractures. *Osteoporosis International*. 2013; Vol. 24, issue Suppl 1:S273-274.

Evstigneeva L, Negodaeva E, Kozhemyakina E, Belkin A. Effect of exercise on balance and functional mobility in postmenopausal women with vertebral fractures. *Osteoporosis International*. 2016; Vol. 27, issue Suppl 1:S520.

Gold 2004 {published data only}

Gold DT, Shipp KM, Pieper CF, Duncan PW, Martinez S, Lyles KW. Group treatment improves trunk strength and psychological status in older women with vertebral fractures: results of a randomized, clinical trial. *Journal of the American Geriatrics Society* 2004;**52**(9):1471-8.

Malmros 1998 {published data only}

Malmros B, Mortensen L, Jensen MB, Charles P. Positive effects of physiotherapy on chronic pain and performance in osteoporosis. *Osteoporosis International* 1998;**8**(3):215-21.

Papaioannou 2003a {published data only}

Papaioannou A, Adachi JD, Winegard K, Ferko N, Parkinson W, Cook RJ, et al. Efficacy of home-based exercise for improving quality of life among elderly women with symptomatic osteoporosis-related vertebral fractures. *Osteoporosis International* 2003;**14**(8):677-82.

Wang 2015 {published data only}

Wang XF, Xu B, Ye XY, Yang YY, Wang GH. Effects of different treatments on patients with osteoporotic fracture after percutaneous kyphoplasty. *Zhongguo Gu Shang* 2015;**28**(6):512-6.

Yang 2007 {published data only}

Yang L, He CQ, Xie W, Lan Q. Effect of pain-free exercises on female osteoporosis patients with spinal compressive fracture. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2007;**11**(45):9108-11.

References to studies excluded from this review

Bada 2009 {published data only}

Bada AM, Rodriguez JE. What are the best treatments for reducing pain due to compression fractures?. *Evidence Based Practice* 2009;**12**(1):4-5.

Barker 2014 {published data only}

Barker KL, Javaid MK, Newman M, Lowe CM, Stallard N, Campbell H, et al. Physiotherapy Rehabilitation for Osteoporotic Vertebral Fracture (PROVE): study protocol for a randomised controlled trial. *Trials* 2014;**15**(22):1-11.

Bautmans 2010 {published data only}

Bautmans I, Van Arken J, Van Mackelenberg M, Mets T. Rehabilitation using manual mobilization for thoracic kyphosis in elderly postmenopausal patients with osteoporosis. *Journal of Rehabilitation Medicine* 2010;**42**:129-35.

Bergland 2012 {published data only}

Bergland A. Effect of exercise on falls efficacy in osteoporotic women with a history of vertebral fracture: A randomized controlled trial. *8th Congress of the European Geriatric Medicine* 2012;**35**:S67.

Borgo 2010 {published data only}

Borgo MJ, Sinaki M. Back progressive resistive exercise program to reduce risk of vertebral fractures. *Journal für Mineralstoffwechsel* 2010;**17**(2):66-71.

Chan 2017 {published data only}

Chan DC, Chang CB, Liao JK, Han DS, Hwang JS, Liu PY, Hong CH, Tsai KS, Huang KC, Yang RS. Effects of two exercise programs on posture and back pain in community-dwelling elderly with high fragility fracture risks. *Osteoporosis International*, 2017, Conference: World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases, WCO-IOF-ESCEO 2017. Italy. 28 (Suppl 1):S127-S636. 2017.

Chen 2011 {published data only}

Chen BL, Zhong Y, Huang YL, Zeng LW, Li YQ, Yang XX, et al. Systematic back muscle exercise after percutaneous vertebroplasty for spinal osteoporotic compression fracture patients: a randomized fracture patients: a randomized controlled trial. *Clinical Rehabilitation* 2011;**26**(6):483-92.

Chen 2015 {published data only}

Chen JP, Qi XW, Li SJ, Kuang LP, Yuan XH, Wang GS, et al. Bone cement injection as vertebral augmentation therapy for osteoporotic vertebral compression fractures. *Journal of Clinical Rehabilitative Tissue Engineering Research* 2015;**19**(21):3292.

Dusdal 2011 {published data only}

Dusdal K, Grundmanis J, Luttin K, Ritchie P, Rompre C, Sidhu R, et al. Effects of therapeutic exercise for persons with osteoporotic vertebral fractures: a systematic review. *Osteoporosis International* 2011;**22**:755-69.

Evstigneyeva 2013 {published data only}

Evstigneyeva L, Kozhemyakina E, Lesnyak O, Negodaeva E, Guselnikova G, Belkin A, et al. Effect of exercise for patients with osteoporotic vertebral fractures. *Annals of the Rheumatic Diseases*. 2013; Vol. 72:A137.

Gandhi 2015 {published data only}

Gandhi V, Cooper C, Barker K. Recruitment challenges in physiotherapy trials. *Trials* 2015 16(Suppl 2):P95. 2015. [DOI: [10.1186/1745-6215-16-S2-P95](https://doi.org/10.1186/1745-6215-16-S2-P95)]

Giangregorio 2014c {published data only}

Giangregorio LM, Thabane L, Adachi JD, Ashe MC, Bleakney RR, Braun EA, et al. Build better bones with exercise: protocol for a feasibility study of a multicenter randomized controlled trial of 12 months of home exercise in women with a vertebral fracture. *Physical Therapy* 2014;**94**(9):1337-52. [NCT01761084]

Hongo 2007 {published data only}

Hongo M, Itoi E, Sinaki M, Miyakoshi N, Shimada Y, Maekawa S, et al. Effect of low-intensity back exercise on quality of life and back extensor strength in patients with osteoporosis: a randomized controlled trial. *Osteoporosis International* 2007;**18**(10):1389-95.

Jensen 2012 {published data only}

Jensen RK, Leboeuf-Yde C, Wedderkopp N, Sorensen JS, Manniche C. Rest versus exercise as a treatment for patients with low back pain and Modic changes. a randomized controlled clinical trial. *BMC Medicine* 2012;**10**(22):1-13.

Kaijser Alin 2016 {published data only}

Kaijser Alin K, Grahns Kronhed AL, Uzunel E, Alinaghizadeh H, Salminen H. Effect of treatment on back extensor strength with a spinal orthosis in elderly women with osteoporosis: a randomized controlled trial. *Osteoporosis International* (2016) 27 (Suppl 1):S39-S62. 2016.

Karakasidou 2013 {published data only}

Karakasidou P, Skordilis EK, Lyrakis GP. Motor control exercise can reduce pain and kyphosis in osteoporotic women with

vertebral fractures: a randomized controlled trial. *Review of Clinical Pharmacology and Pharmacokinetics* 2013;**27**:95-106.

Liu 2017 {published data only}

Liu T, Sheng L, Huang Z. Influence of exercise intervention on pain, quality of life and functional mobility in patients with osteoporotic vertebral fractures: a systematic review and meta-analysis. *International Journal of Clinical and Experimental Medicine* 2017;**10**(7):9864-70.

Lord 1996 {published data only}

Lord SR, Ward JA, Williams P, Zivanovic E. The effects of a community exercise program on fracture risk factors in older women. *Osteoporosis International* 1996;**6**(5):361-7.

Malmros 1999 {published data only}

Malmros B, Mortensen L, Jensen M, Charles P. Effect of specific physiotherapy on chronic pain, functional level and quality of life in osteoporosis: A prospective randomized single-blind placebo controlled study. *Ugeskrift Laeger* 1999;**161**(33):4636-41.

Marcu 2015 {published data only}

Marcu RI, Patru S, Bighea AC, Traistaru R, Popescu RS. Role of physical exercise program in patients with osteoporosis. *Annals of the Rheumatic Diseases* 74:529. 2015.

Mihailov 2011 {published data only}

Mihailov M, Popa D, Suci N. Efficacy of a rehabilitation program on pain, daily activities and quality of life in women with osteoporosis. *Osteoporosis International* 2011;**22**(Suppl 1):S295.

Miyakoshi 2010 {published data only}

Miyakoshi N, Horikawa A, Hongo M, Kasukawa Y, Shimada Y. Effects of a simple, low-intensity back extension exercise on quality of life and back extensor strength in patients with osteoporosis and vertebral fractures: a 6-month randomized controlled trial. *Journal of Bone and Mineral Research*. 2010; Vol. Abstract.

Nemes 2014 {published data only}

Nemes D, Catan L, Popa D, Amaricai E, Andrei D, Puenea G. Role of rehabilitation in preventing radiculopathies due to osteoporotic vertebral compression fractures. *Osteoporosis International*. 2014; Vol. 25 Suppl 2:S205.

Rapp 2011 {published data only}

Rapp SM. Early mobilization, rigid fixation after vertebral compression fractures reduces vertebral deformity. *Orthopedics Today* 2011;**October**:45.

Rittchen 1991 {published data only}

Rittchen A, Petravic B. [Fruhfunktionelle Behandlung der stabilen Kompressionsbrüche der Brust- und Lendenwirbelsäule]. *Krankengymnastik* 1991;**43**(4):336-9.

Schwinning 1992 {published data only}

Schwinning M. [Krankengymnastik nach Wirbelfrakturen der BWS-LWS ohne neurologische Symptomatik]. *Krankengymnastik* 1992;**44**(1):34-8.

Shipp 2004 {published data only}

Shipp KM, Pieper, CF, Gold DT, Lyles KW. Change in walking endurance among women with prevalent vertebral fractures: associations with concurrent change in impairment, functional status, and disability. *Journal of Bone and Mineral Research*. 2004; Vol. Abstract M454.

Shipp 2007 {published data only}

Shipp KM, Gold DT, Pieper, CF, Lyles KW. Improving trunk strength and endurance in older women with vertebral fractures. *Journal of Bone and Mineral Research*. 2007; Vol. Abstract W398.

Sinaki 1984 {published data only}

Sinaki M, Mikkelsen BA. Postmenopausal spinal osteoporosis: flexion versus extension exercises. *Archives of Physical Medicine and Rehabilitation* 1984;**65**(10):593-6.

Sinaki 1995 {published data only}

Sinaki M. Rehabilitation of osteoporotic fractures of the spine. *Physical Medicine and Rehabilitation* 1995;**9**(1):105-23.

Smith 1998 {published data only}

Smith AA, Summers GD, Baxendale A, Butterly RJ. A randomized controlled trial of the effects of weight-bearing exercise and resistance training on risk factors for fracture in women with osteoporosis. *Current Research in Osteoporosis and Bone Mineral Measurement*. London: British Institute of Radiology, 1998:128.

Stadhouder 2009 {published data only}

Stadhouder A, Buskens E, Vergroesen DA, Fidler MW, de Nies F, Oner FC. Nonoperative treatment of thoracic and lumbar spine fractures: a prospective randomized study of different treatment options. *Journal of Orthopaedic Trauma* 2009;**23**(8):588-94.

Svensson 2017 {published data only}

Svensson Hk, Olsson LE, Hansson T, Karlsson J, Hansson-Olofsson E. The effects of person-centered or other supportive interventions in older women with osteoporotic vertebral compression fractures—a systematic review of the literature. *Osteoporosis International* 2017;**28**(9):2521-40.

Webber 2003 {published data only}

Webber CE, Papaioannou A, Winegard KJ, Adachi JD, Parkinson W, Ferko NC, et al. A 6-mo home-based exercise program may slow vertebral height loss. *Journal of Clinical Densitometry* 2003;**6**(4):391-400.

Zhong 2012 {published data only}

Zhong Y, Huang Y, Li Y, Wang C. Medium and long term effects of systematic back muscle exercise after percutaneous vertebroplasty for osteoporotic fracture. *Chinese Journal of Tissue Engineering Research* 2012;**16**(30):5544-5548.

References to studies awaiting assessment
Bergstrom 2016 {published data only}

Bergstrom I. Vertebral fractures in an osteoporotic perspective, symptoms and what to do. *Osteoporosis International*. 2016; Vol. 27 Suppl 1:S570.

Giangregorio 2017 {published data only}

Giangregorio LM, Thabane L, Adachi JD, Ashe MC, Bleakney R, Cheung AM, et al. Build better bones with exercise: feasibility of a multicenter randomized controlled trial of 12 months of home exercise in women with a vertebral fracture. *Journal of Bone and Mineral Research*. 2017; Vol. 32 (Suppl 1). [NCT01761084]

Giangregorio 2018 {published data only}

* Giangregorio LM, Gibbs JC, Templeton JA, Adachi JD, Ashe MC, Bleakney RR, et al. Build better bones with exercise (B3E pilot trial): results of a feasibility study of a multicenter randomized controlled trial of 12 months of home exercise in older women with vertebral fracture. *Osteoporosis International* 2018;**29**:2545-56. [DOI: [10.1007/s00198-018-4652-0](https://doi.org/10.1007/s00198-018-4652-0)]

Pompa 2016 {published data only}

Pompa A, Lugari G, Giampa E, Della Vedova C, Pace L, Tramontano M, et al. Effects of postural and respiratory rehabilitation in patients with hyper-kypnosis by osteoporotic vertebral fractures: a pilot study. *Osteoporosis International* 2016;**27**(Suppl 1):S122.

References to ongoing studies
Stanghelle 2018 {published data only}

NCT02781974. The effect of exercise on physical function in older women with osteoporosis and a history of vertebral fracture: a randomized controlled trial. clinicaltrials.gov/show/NCT02781974 (first posted May 25, 2016). [ClinicalTrials.gov: NCT02781974]

* Stanghelle B, Bentzen H, Giangregorio L, Pripp AH, Bergland A. Effect of a resistance and balance exercise programme for women with osteoporosis and vertebral fracture: study protocol for a randomized controlled trial. *BMC Musculoskeletal Disorders* 2018;**19**:100. [DOI: [10.1186/s12891-018-2021-y](https://doi.org/10.1186/s12891-018-2021-y)]

Additional references
Adachi 2001

Adachi JD, Loannidis G, Berger C, Joseph L, Papaioannou A, Pickard L, et al. The influence of osteoporotic fractures on health-related quality of life in community-dwelling men and women across Canada. *Osteoporosis International* 2001;**12**(11):903-8.

Avenell 2014

Avenell A, Mak JC, O'Connell D. Vitamin D and vitamin D analogues for preventing fractures in post-menopausal women and older men. *Cochrane Database of Systematic Reviews* 2014, Issue 4. [DOI: [10.1002/14651858.CD000227.pub4](https://doi.org/10.1002/14651858.CD000227.pub4)]

Bessette 2008

Bessette L, Ste-Marie LG, Jean S, Davison KS, Beaulieu M, Baranci M, et al. Recognizing osteoporosis and its consequences in Quebec (ROCQ); background, rationale and methods of an anti-fracture patient health-management programme. *Contemporary Clinical Trials* 2008;**29**(2):194-210.

Black 1996

Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. *Lancet* 1996 Dec 7;**348**(9041):1535-41.

Briggs 2007

Briggs AM, van Dieën JH, Wrigley TV, Greig AM, Phillips B, Lo SK, et al. Thoracic kyphosis affects spinal loads and trunk muscle force. *Physical Therapy* 2007;**87**(5):595-607.

Bruno 2012

Bruno AG, Anderson DE, D'Agostino J, Bouxsein ML. The effect of thoracic kyphosis and sagittal plane alignment on vertebral compressive loading. *Journal of Bone and Mineral Research* 2012;**27**(10):2144-51.

Cauley 2000

Cauley JA, Thompson DE, Ensrud KC, Scott JC, Black D. Risk of mortality following clinical fractures. *Osteoporosis International* 2000;**11**(7):556-61.

Cauley 2007

Cauley JA, Hochberg MC, Lui LY, Palermo L, Ensrud KE, Hillier TA, et al. Long-term risk of incident vertebral fractures. *JAMA* 2007;**298**(23):2761-7.

Chou 2012

Chou CH, Hwang CL, Wu YT. Effect of exercise on physical function, daily living activities, and quality of life in the frail older adults: a meta-analysis. *Archives of Physical Medicine and Rehabilitation* 2012;**93**(2):237-44.

Cooper 1993

Cooper C, O'Neill T, Silman A. The epidemiology of vertebral fractures. European Vertebral Osteoporosis Study Group. *Bone* 1993;**14**(Suppl 1):S89-97.

Dusdal 2010

Dusdal K, Grundmanis J, Luttin K, Ritchie P, Rompre C, Sidhu R, et al. Effects of therapeutic exercise for persons with osteoporotic vertebral fractures: a systematic review. *Osteoporosis International* 2010;**22**:755-69.

Evstigneeva 2013

Evstigneeva L, Lesnyak O, Negodaeva E, Guselnikova G, Kozhemyakina E. A randomized, prospective study of the effects of exercise on quality of life in postmenopausal women with vertebral fractures. *Osteoporosis International*. 2013; Vol. 24(Suppl 1):S273-274.

FDA 1995

Expert Working Group (Efficacy) of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Clinical Safety Data Management: Definitions and Standards for Expedited Reporting. FDA Center for Drug and Evaluation Research 1995.

Gallacher 2007

Gallacher SJ, Gallagher AP, McQuillian C, Mitchell PJ, Dixon T. The prevalence of vertebral fracture amongst patients presenting with non-vertebral fractures. *Osteoporosis International* 2007;**18**(2):185-92.

Gautschi 2017

Gautschi OP, Stienen MN, Corniola MV, Joswig H, Schaller K, Hildebrandt G, et al. Assessment of the Minimum Clinically Important Difference in the Timed Up and Go Test After Surgery for Lumbar Degenerative Disc Disease. *Neurosurgery* 2017;**80**(3):380-5.

Ghogomu 2014

Ghogomu EA, Maxwell LJ, Buchbinder R, Rader T, Pardo Pardo J, Johnston RV, et al. Editorial Board of the Cochrane Musculoskeletal Group. Updated method guidelines for Cochrane Musculoskeletal Group systematic reviews and metaanalyses. *Journal of Rheumatology* 2014 Feb;**41**(2):194-205. [DOI: [10.3899/jrheum.121306](https://doi.org/10.3899/jrheum.121306)]

Giangregorio 2014a

Giangregorio LM, Papaioannou A, Macintyre NJ, Ashe MC, Heinonen A, Shipp K, et al. Too Fit to Fracture: exercise recommendations for individuals with osteoporosis or osteoporotic vertebral fracture. *Osteoporosis International* 2014 March;**25**(3):821-35.

Giangregorio 2014b

Giangregorio LM, MacIntyre NJ, Heinonen A, Cheung AM, Wark JD, Shipp K, et al. Too Fit To Fracture: a consensus on future research priorities in osteoporosis and exercise. *Osteoporosis International* 2014 May;**25**(5):1465-72.

Giangregorio 2015

Giangregorio LM, McGill S, Wark JD, Laprade J, Heinonen A, Ashe MC, et al. Too Fit To Fracture: outcomes of a Delphi consensus process on physical activity and exercise recommendations for adults with osteoporosis with or without vertebral fractures. *Osteoporosis International* 2015 March;**26**(3):891-910.

Handoll 2011

Handoll HH, Sherrington C, Mak JC. Interventions for improving mobility after hip fracture surgery in adults. *Cochrane Database of Systematic Reviews* 2011, Issue 3. [DOI: [10.1002/14651858.CD001704.pub4](https://doi.org/10.1002/14651858.CD001704.pub4)]

Handoll 2015

Handoll HH, Elliott J. Rehabilitation for distal radial fractures in adults. *Cochrane Database of Systematic Reviews* 2015, Issue 9. [DOI: [10.1002/14651858.CD003324.pub3](https://doi.org/10.1002/14651858.CD003324.pub3)]

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Higgins 2011

Higgins JP, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org. The Cochrane Collaboration.

Howe 2011

Howe TE, Shea B, Dawson LJ, Downie F, Murray A, Ross C, et al. Exercise for preventing and treating osteoporosis in postmenopausal women. *Cochrane Database of Systematic Reviews* 2011, Issue 7. [DOI: [10.1002/14651858.CD000333.pub2](https://doi.org/10.1002/14651858.CD000333.pub2)]

Jackson 2000

Jackson SA, Tenenhouse A, Robertson L. Vertebral fracture definition from population-based data: preliminary results from the Canadian Multicenter Osteoporosis Study (CaMos). *Osteoporosis International* 2000;**11**(8):680-7.

Kado 2007

Kado DM, Huang MH, Nguyen CB, Barrett-Connor E, Greendale GA. Hyperkyphotic posture and risk of injurious falls in older persons: the Rancho Bernardo Study. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences* 2007;**62**(6):652-7.

Kanis 2001

Kanis JA, Oden A, Johnell O, Jonsson B, De Laet C, Dawson A. The burden of osteoporotic fractures: a method for setting intervention thresholds. *Osteoporosis International* 2001;**12**(5):417-27.

Kanis 2004

Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone* 2004;**35**(2):375-82.

Kastner 2007

Kastner M, Straus S, Goldsmith CH. Estimating the Horizon of articles to decide when to stop searching in systematic reviews: an example using a systematic review of RCTs evaluating osteoporosis clinical decision support tools. *AMIA Annual Symposium Proceedings*. 2007:389-93.

Kastner 2009

Kastner M, Straus SE, McKibbin KA, Goldsmith CH. The capture-mark-recapture technique can be used as a stopping rule when searching in systematic reviews. *Journal of Clinical Epidemiology* 2009;**62**(2):149-57.

Kelley 1998

Kelley GA. Exercise and regional bone mineral density in postmenopausal women: a meta-analytic review of randomized trials. *American Journal of Physical Medicine and Rehabilitation* 1998;**77**(1):76-87.

Lips 2005

Lips P, van Schoor NM. Quality of life in patients with osteoporosis. *Osteoporosis International* 2005 May;**16**(5):447-55.

Little 1987

Little RJ, Rubin DB. *Statistical Analysis with Missing Data*. 2nd Edition. New York: John Wiley and Sons, 1987.

Martyn-St James 2009

Martyn-St James M, Carroll S. A meta-analysis of impact exercise on postmenopausal bone loss: the case for mixed loading exercise programmes. *British Journal of Sports Medicine* 2009;**43**(12):898-908.

Moayyeri 2008

Moayyeri A. The association between physical activity and osteoporotic fractures: a review of the evidence and implications for future research. *Annals of Epidemiology* 2008;**18**(11):827-35.

Oleksik 2000

Oleksik A, Lips P, Dawson A, Minshall ME, Shen W, Cooper C, et al. Health-related quality of life in postmenopausal women with low BMD with or without prevalent vertebral fractures. *Journal of Bone and Mineral Research* 2000;**15**(7):1384-92.

Olsen 2014

Olsen CF, Bergland A. The effect of exercise and education on fear of falling in elderly women with osteoporosis and a history of vertebral fracture: results of a randomized controlled trial. *Osteoporosis International* 2014;**25**(8):2017-25.

Pahor 2014

Pahor M, Guralnik JM, Ambrosius WT, Blair S, Bonds DE, Church TS, et al. LIFE study investigators. Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. *JAMA* 2014 Jun 18;**311**(23):2387-96.

Papaioannou 2002

Papaioannou A, Watts NB, Kendler DL, Yuen CK, Adachi JD, Ferko N. Diagnosis and management of vertebral fractures in elderly adults. *American Journal of Medicine* 2002;**113**(3):220-8.

Papaioannou 2003b

Papaioannou A, Parkinson W, Ferko N, Probyn L, Ioannidis G, Jurriaans E, et al. Prevalence of vertebral fractures among patients with chronic obstructive pulmonary disease in Canada. *Osteoporosis International* 2003;**14**(11):913-7.

Petrella 2000

Petrella RJ, Payne M, Myers A, Overend T, Chesworth B. Physical function and fear of falling after hip fracture rehabilitation in the elderly. *American Journal of Physical Medicine and Rehabilitation* 2000;**79**(2):154-60.

Polidoulis 2012

Polidoulis I, Beyene J, Cheung AM. The effect of exercise on pQCT parameters of bone structure and strength in postmenopausal women--a systematic review and meta-

analysis of randomized controlled trials. *Osteoporosis International* 2012 Jan;**23**(1):39-51.

Salaffi 2004

Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *European Journal of Pain* 2004 Aug;**8**(4):283-91.

Schousboe 2008

Schousboe JT, Vokes T, Broy SB, Ferrar L, McKiernan F, Roux C, et al. Vertebral Fracture Assessment: the 2007 ISCD Official Positions. *Journal of Clinical Densitometry* 2008;**11**(1):92-108.

Sherrington 2011

Sherrington C, Tiedemann A, Fairhall N, Close J, Lord SR. Exercise for preventing falls in older adults: an updated meta-analysis and best practice recommendations. *NSW Public Health Bulletin* 2011;**22**(3-4):78-83.

Wells 2008a

Wells GA, Cranney A, Peterson J, Boucher M, Shea B, Welch V, et al. Alendronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: [10.1002/14651858.CD001155.pub2](https://doi.org/10.1002/14651858.CD001155.pub2)]

Wells 2008b

Wells GA, Cranney A, Peterson J, Boucher M, Shea B, Welch V, et al. Etidronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: [10.1002/14651858.CD003376.pub3](https://doi.org/10.1002/14651858.CD003376.pub3)]

Wells 2008c

Wells GA, Cranney A, Peterson J, Boucher M, Shea B, Welch V, et al. Risedronate for the primary and secondary prevention of osteoporotic fractures in postmenopausal women. *Cochrane Database of Systematic Reviews* 2008, Issue 1. [DOI: [10.1002/14651858.CD004523.pub3](https://doi.org/10.1002/14651858.CD004523.pub3)]

WHO 2003

WHO Scientific Group on the Prevention and Management of Osteoporosis. Prevention and management of osteoporosis: report of a WHO scientific group. WHO Technical Report Series 2003; Vol. 921.

Wiktorowicz 2001

Wiktorowicz ME, Goeree R, Papaioannou A, Adachi JD, Papadimitropoulos E. Economic implications of hip fracture: health service use, institutional care and cost in Canada. *Osteoporosis International* 2001;**12**(4):271-8.

Wright 2011

Wright AA, Cook CE, Baxter GD, Dockerty JD, Abbott JH. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *Journal of Orthopaedic and Sports Physical Therapy* 2011;**41**(5):319-27.

References to other published versions of this review

Giangregorio 2010

Giangregorio LM, MacIntyre N, Thabane L, Cranney A, Cook R, Papaioannou A. Exercise for improving outcomes after osteoporotic vertebral fracture. *Cochrane Database of Systematic Reviews* 2010, Issue 7. [DOI: [10.1002/14651858.CD008618](https://doi.org/10.1002/14651858.CD008618)]

Giangregorio 2013

Giangregorio LM, Macintyre NJ, Thabane L, Skidmore CJ, Papaioannou A. Exercise for improving outcomes after osteoporotic vertebral fracture. *Cochrane Database of Systematic Reviews* 2013, Issue 1. [DOI: [10.1002/14651858.CD008618.pub2](https://doi.org/10.1002/14651858.CD008618.pub2)]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies [ordered by study ID]

Bennell 2010

Methods	Randomized controlled trial
Participants	# Randomized: 20 (11 intervention, 9 control) Lost to follow-up: 0 Age: 53-90 (66.2 intervention, 66.3 control) Country: Australia Inclusion criteria: at least 1 painful vertebral crush or wedge fracture ($\geq 20\%$ reduction in anterior height relative to posterior) sustained 3-24 months previously, osteoporosis (lumbar spine or proximal femur BMD T-score ≤ -2.5), 5 years postmenopause, stable medications for 6 months, English speaking, community dwelling.

Bennell 2010 (Continued)

Exclusion criteria: contraindications to exercise (heart or pulmonary disease, inflammatory or joint disease or severe osteoarthritis, psychiatric condition, neuromuscular condition), secondary causes of osteoporosis e.g. osteomalacia, radicular signs or symptoms, back pain radiating into lower limb, physiotherapy in the last 6 months or previous participation in pain management program for back pain, allergy to adhesive tape or skin condition preventing use of tape.

Interventions
Experimental

Frequency: strengthening 3 times/week, physical therapy once per week, posture training and range of motion exercises daily

Intensity: varies according to clinical presentation

Time: not specified

Type: multi-modal physical therapy intervention including manual therapy, massage, taping, posture training and exercise prescription.

Duration of intervention: 10 weeks

Supervisor: physical therapist

Supervision/Setting: intermittent, clinic and home

Progression of intensity: unknown

Control: no intervention

Duration of follow-up after baseline assessment: 11 weeks

Outcomes

Pain: 11-point rating scale for pain at rest and pain with activities

Quality of life: QUALEFFO-41 (disease-specific), assessment of Quality of Life

Physical performance: Timed Up and Go, 11-point scale for restriction of daily activities

Posture: Dualer Electric Inclinometer

Trunk and arm endurance: Timed Loaded Standing

Notes

Adherence: 82%

Adverse events

Due to intervention:

- shoulder pain (2, 10%)

- flare up of wrist injury

- sore wrist

- sore knee

- tape irritation

Unclear if due to intervention:

- none

Not due to intervention:

- none

Bennell 2010 (Continued)

Between-group comparison as actual difference or change scores: actual difference adjusted for baseline values (ANCOVA)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The randomisation sequence was generated a priori using the random number function in Excel (Microsoft Corporation, USA) by an independent investigator not directly involved in assessment of participants. These were concealed in opaque envelopes and stored in a locked central location." Comment: probably done.
Allocation concealment (selection bias)	Low risk	Quote: "The envelopes were opened sequentially at the time of intervention assignment by an administrator who informed the treating therapist by facsimile or email." Comment: probably done.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blinded to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "Participants were assessed at baseline and at 11 weeks by an assessor blinded to group allocation." Comment: probably done.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: pain and quality of life were reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "All analyses were conducted on an intention-to-treat principle using all randomized participants using the Statistical Package for the Social Sciences (Norusis/ SPSS Inc., Chicago IL, USA)." Comment: probably done. All participants completed the trial.
Selective reporting (reporting bias)	Unclear risk	Comment: the clinical trial registration number was reported, and upon reviewing the registered trial, two outcomes were listed that were not reported in the published article: standing balance on a force platform, and Human Activity Profile. Communication with the authors revealed that they were not able to process the standing balance data, so the choice to omit the data was due to logistical difficulties. The Human Activity Profile Data was collected for descriptive purposes.

Bergland 2011

Methods	Randomized controlled trial
Participants	# Randomized: 89 (47 intervention, 42 control) Lost to follow-up: 11 (6 intervention, 5 control) at 3 months, 6 additional (1 intervention, 5 control) at 12 months Age: 71.4 (70.8 intervention, 72 control)

Bergland 2011 (Continued)

Country: Norway

Inclusion criteria: living at home, ambulatory, 60 years and older, and at least one vertebral fracture verified by x-ray, osteoporosis (lumbar spine or proximal femur BMD T-score ≤ -2.5).

Exclusion criteria: recent vertebral fractures, unable to complete questionnaires, Mini-Mental State Examination ≤ 23 .

Interventions	<p>Experimental</p> <p><i>Frequency:</i> 2 times a week</p> <p><i>Intensity:</i> adjust to ability</p> <p><i>Time:</i> 60 minutes</p> <p><i>Type:</i> balance training, agility training, posture exercises</p> <p><i>Duration of intervention:</i> 12 weeks</p> <p><i>Supervisor:</i> physiotherapist</p> <p><i>Supervision/Setting:</i> continual supervision in a group exercise class</p> <p><i>Progression of intensity:</i> yes</p> <p>Control: usual activities</p> <p>Duration of follow-up after baseline assessment: 12 weeks and 52 weeks</p>
Outcomes	<p>Physical performance: Timed Up and Go, maximum walking speed over 20 meters in seconds</p> <p>Balance: Functional Reach test</p> <p>Quality of life: QUALEFFO-41, General Health Questionnaire</p> <p>Outcomes published in Olsen 2014 only:</p> <p>Fear of falling: Falls Self-Efficacy Scale-I</p> <p>Self-reported falls during study period</p>
Notes	<p>Adherence: 81%</p> <p>Adverse events: none reported</p> <p>Between-group comparison as actual difference or change scores: change scores</p> <p>Note: Olsen 2014 published a secondary analysis of new data on the effect of exercise on self-reported number of falls and fear of falling in participants, which were not reported in the original trial by Bergland 2011.</p>

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	<p>Quote: "Participants were randomly assigned by a computer-generated list to two groups, intervention (IT) and control (CT)."</p> <p>Comment: probably done.</p>
Allocation concealment (selection bias)	Low risk	<p>Quote: "Researchers not involved in the study performed the randomization by using lots in sealed opaque envelopes."</p>

Bergland 2011 (Continued)

		Comment: probably done
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "All the subjects included in the study were assessed three times by a physiotherapist who did not know which of the two groups the participants belonged to." Comment: probably done.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: falls, quality of life questionnaires, and fear of falling were reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Quote: "Analysis was on an intention-to-treat basis, in which missing data at 3 or 12 months for subjects who did not complete the program were replaced by baseline test values." Comment: probably done.
Selective reporting (reporting bias)	Unclear risk	Comment: the protocol is not publicly available for review or the clinical trial registration number was not provided. Olsen 2014 published a secondary analysis of new data on the effect of exercise on self-reported number of falls and fear of falling in participants, which were not reported in the original trial by Bergland 2011 .

Bergstrom 2011

Methods	Randomized controlled trial
Participants	<p># Randomized: 36 (20 intervention, 16 control)</p> <p>Lost to follow-up: 3 (2 intervention, 1 control) at 4 months</p> <p>Age: 73.2 intervention, 74.1 control</p> <p>Country: Sweden</p> <p>Inclusion criteria: at least one vertebral fracture (at least 6 months prior), osteoporosis (states diagnosis by DXA but no criteria), bisphosphonate therapy for 4 months, kyphosis (no criteria stated), postmenopausal (no time criteria stated)</p> <p>Exclusion criteria: metabolic disease or secondary osteoporosis, cardiovascular or musculoskeletal complications that would prevent full participation, already participating in exercise at same level</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> 2 times a week</p> <p><i>Intensity:</i> individualized to ability, 30 repetitions of each exercise</p> <p><i>Time:</i> 60 minutes</p> <p><i>Type:</i> warm up, exercises for back extensors, shoulder flexors and shoulder abductors, scapular retraction and push-ups, bird dog (thoracic and lumbar extension)</p> <p><i>Duration of intervention:</i> 16 weeks</p>

Bergstrom 2011 (Continued)

Supervisor: physiotherapist

Supervision/Setting: continual supervision in a group exercise class

Progression of intensity: yes

Control: not described

Duration of follow-up after baseline assessment: 16 weeks

Outcomes	Back Extension Strength: in standing with an isometric dynamometer (in Newtons) Height Posture - kyphometer and C7 to wall distance
Notes	Adherence: not specifically stated. Per-protocol analyses included those with at least 90% attendance to sessions, two were excluded because of poor (< 50%) attendance, and two were lost to follow-up. Although goal was strengthening, 30 repetitions were performed, suggesting that endurance would be the goal.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Comment: participants picked allocation out of hat.
Allocation concealment (selection bias)	High risk	Comment: study nurse prepared materials for participant randomization.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "All measurements were ...blinded as to the participant. The study nurse examined and evaluated all women." Comment: it is not clear if the study nurse that performed randomization also performed outcome assessment.
Blinding of subjective outcome assessment (detection bias)	Unclear risk	Comment: it is not clear if the study nurse that performed randomization also performed outcome assessment.
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: per-protocol analysis performed only for all secondary outcomes. Although intention-to-treat analysis was performed for the primary outcome, per-protocol analysis was also performed in two ways (with and without adjustment for baseline levels of outcome) and the per-protocol analysis is emphasized in interpretation.
Selective reporting (reporting bias)	Unclear risk	Comment: no clinical trial registration number available.

Evstigneeva 2016

Methods	Randomized controlled trial
Participants	<p># Randomized: 78 (40 intervention, 38 control)</p> <p>Lost to follow-up: 2 (1 intervention, 1 control); 2 did not receive allocated intervention</p> <p>Age: 70.7 intervention, 67.6 control</p> <p>Country: Russia</p> <p>Inclusion criteria: at least one radiographically confirmed vertebral fracture, chronic back pain, post-menopausal osteoporosis, aged 50 years or older</p> <p>Exclusion criteria: vertebral and non vertebral fracture within the past 6 months, secondary osteoporosis, osteomalacia, inflammatory joint or spine disease, diffuse connective tissue disease, history of malignancy in the last 5 years, significant functional abnormalities of the cardiovascular, respiratory, gastrointestinal, or nervous system, renal disease, and corticosteroid intake during the last year</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> 2 times a week</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> 40 minutes</p> <p><i>Type:</i> physical exercise program including strength training for back extensors, upper and lower extremities, and abdomen, weight-bearing exercises, relaxation techniques and autogenic training</p> <p><i>Duration of intervention:</i> 52 weeks</p> <p><i>Supervisor:</i> instructor- type not specified</p> <p><i>Supervision/Setting:</i> continual supervision at a center-based facility with a 1:1 ratio</p> <p><i>Progression of intensity:</i> yes</p> <p>Note: if regular attendance to center-based exercise sessions was impossible, participants were allowed to perform the structured exercise program at home with properly training from the instructor, and a booklet and DVD disk describing exercise program was provided.</p> <p>Control: usual activities</p> <p>Duration of follow-up after baseline assessment: 52 weeks</p>
Outcomes	<p>Quality of life (disease-specific): QUALEFFO-41</p> <p>Dynamic balance: Tandem walk and sway, sit-to-stand weight transfer and left/right weight symmetry, weight-bearing squat test</p> <p>Physical performance: Timed Up and Go test</p> <p>Posture: occiput-to-wall distance</p> <p>Incident fractures</p>
Notes	<p>Adherence: 89.2%</p> <p>Adverse events: one participant reported worsening knee pain</p> <p>Between-group comparison as actual difference or change scores: change scores</p>

Risk of bias

Evstigneeva 2016 (Continued)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "Patients were randomly assigned to the exercise and control groups in a 1:1 ratio." Comment: probably done.
Allocation concealment (selection bias)	Low risk	Quote: "Randomization was performed using pre-generated randomization numbers in sealed opaque envelopes. An independent person not having any information about the patients opened the envelopes. Randomization number was reported to a researcher by telephone." Comment: probably done.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Quote: "Blinded assessment of the parameters was conducted in the implementation and evaluation of stabilimetry and quality of life questionnaire." Comment: blinded assessment was conducted in implementation and evaluation of stabilimetry and QOL questionnaire; however, it was unclear if occiput-to-wall distance and Timed Up and Go test assessments were blinded.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: disease-specific quality of life was reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Quote: "Data were analyzed following an intention-to-treat principle and missing data were managed using the last observation carried forward (LOCF) technique." Comment: two participants did not receive the allocated intervention and it was unclear if they were included in the analysis and why they did not perform the exercises.
Selective reporting (reporting bias)	Unclear risk	Two abstracts by Evstigneeva 2013 and Evstigneyeva 2013 reported on a subset of outcomes (QUALEFFO-41, Timed Up and Go test, and stabilimetry- Sit-to-Stand Weight Transfer, Body Weight Rising Index, and Tandem Walk and Sway) from the larger database included in Evstigneeva 2016 . The stabilimetry balance outcomes reported on differed slightly between Evstigneyeva 2013 and Evstigneeva 2016 . All three studies reported Timed Up and Go test results.

Gold 2004

Methods	Cluster-randomized controlled trial (unit of randomization was retirement facility)
Participants	# Randomized: 185 (94 intervention, 91 control) Lost to follow-up: 9 intervention, 5 control Age: 81 Country: USA

Gold 2004 (Continued)

Inclusion criteria: women, at least one vertebral fracture (reduction in vertebral height compared to adjacent vertebrae of 20% or greater)

Exclusion criteria: vertebral fracture in the last 6 months, hip fracture in the last 12 months, more than 2 errors on the short portable mental status questionnaire, corrected vision worse than 20/40, active cardiac, pulmonary, neurological disease or rheumatoid arthritis, injurious fall in the last 6 months, not an independent ambulator

Interventions	<p>Experimental</p> <p><i>Frequency:</i> 3 times/week for exercise, 2 times/week for coping class</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> 45 minutes each for exercise class and coping class</p> <p><i>Type:</i> stretching, strengthening with weights and resistance bands, instruction on optimal alignment and body mechanics, coping class</p> <p><i>Duration of intervention:</i> 6 months</p> <p><i>Supervisor:</i> physical therapist</p> <p><i>Supervision/Setting:</i> group class in retirement facility</p> <p><i>Progression of intensity:</i> yes</p> <p>Control: weekly class on general health information</p> <p>Duration of follow-up after baseline assessment: 6 months</p>
Outcomes	<p>Trunk extension strength: peak isometric strength in foot pounds- mean of 3 maximal trials using B-200 isostation</p> <p>Pain - subscale of Functional Status Index</p> <p>Psychological symptoms - Global Severity Index of the Hopkins Symptom Checklist 90 - Revised</p>
Notes	<p>Adherence: 58%</p> <p>Adverse events</p> <p><i>Due to intervention:</i></p> <ul style="list-style-type: none"> - fracture of costal cartilage while rolling from supine to prone - pain <p><i>Unclear if due to intervention:</i></p> <ul style="list-style-type: none"> - hip fracture when lifting leg into measurement device (6-month assessment) - metatarsal fracture when weight dropped on foot <p><i>Not due to intervention:</i></p> <ul style="list-style-type: none"> - none <p>Between-group comparison as actual difference or change scores: change scores</p>
Risk of bias	
Bias	<p>Authors' judgement</p> <p>Support for judgement</p>

Gold 2004 (Continued)

Random sequence generation (selection bias)	Low risk	Quote: "Randomization was performed using a random number generator with equal allocation to the two arms." Comment: probably done.
Allocation concealment (selection bias)	Low risk	Quote: "All researchers except the biostatistician were masked to allocation status until a site was enrolled." Comment: probably done.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Quote: "Participants were unaware of the content of the intervention that they did not receive initially." Comment: probably done.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "All personnel involved with subject contacts, data collection, and intervention administration were masked to the intervention status of the sites and to the study hypotheses throughout the trial." Comment: probably done.
Blinding of subjective outcome assessment (detection bias)	Unclear risk	Comment: Pain and Global Severity Index were reported by participants who were unaware of group allocation.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Comment: data were excluded from 63 of a total of 185 participants for the muscle strength outcome because of equipment problems; analyses for this outcome used a sample size of 122.
Selective reporting (reporting bias)	Unclear risk	Comment: Gold 2004 reported on data from three primary outcomes. However, the data belong to a much larger database that includes data from a number of studies, and a number of other outcomes were collected but were not reported in the paper. Data on walking endurance (six-minute walk distance) was presented at the American Society for Bone and Mineral Research Annual Meeting in 2004, and data on trunk endurance (Timed Loaded Standing) were presented at the American Society for Bone and Mineral Research Annual Meeting in 2007. We have been in contact with the authors to get access to these data and other unreported outcomes, but they were not available in time for this review. We were informed that the three primary outcomes were chosen a priori.

Malmros 1998

Methods	Randomized controlled trial
Participants	# Randomized: 53 Lost to follow-up: 3 intervention, 2 control, no ITT analysis Age: 65 intervention, 68 control Country: Denmark Inclusion criteria: women, at least 1 vertebral fracture, 55-75 years of age, back pain in the last 5 years (not from a new fracture) Exclusion criteria: no acute back pain from a new fracture, severe heart or pulmonary disease, psychiatric disease, medicine or alcohol abuse, rheumatoid disease

Malmros 1998 (Continued)

Interventions	Experimental <i>Frequency:</i> 2 times/week <i>Intensity:</i> not specified <i>Time:</i> 1 hour <i>Type:</i> balance training, muscle strengthening, including flexors and extensors of trunk, stretching <i>Duration of intervention:</i> 10 weeks <i>Supervisor:</i> physical therapist <i>Supervision/Setting:</i> continual 1:1, center-based <i>Progression of intensity:</i> yes Control: usual activities Duration of follow-up after baseline assessment: weekly for 22 weeks	
Outcomes	Pain: 11-point scale Self-reported physical function: modified Oswestry questionnaire Muscle function: back extension strength in kg measured with strain gauge Quality of life: self-designed Balance: postural sway in x and y with eyes open and closed, tandem and with platform tilted, sum of four scores	
Notes	Adherence: 100% (except for those excluded because of adverse events) Adverse events: 5 events of unknown etiology described as not due to intervention: 3 in intervention group and 2 in control group Between-group comparison as actual difference or change scores: both	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Quote: "The participants were randomized consecutively and in masked fashion by the drawing of sealed envelopes containing the name of the group." Comment: probably not random sequence generation.
Allocation concealment (selection bias)	High risk	Comment: see above. It is unclear who prepared or had access to the envelopes or performed randomization.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "The tester was the same at all visits and was masked with regard to the group to which the participant was allocated." Comment: probably done.

Malmros 1998 (Continued)

Blinding of subjective outcome assessment (detection bias)	High risk	Comment: pain, self-reported physical function and quality of life were reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: the original study would be assessed with a high risk of bias because an as-treated analysis was performed. However, our analysis presented in this review represents an ITT analysis, so the risk of bias in the findings we report is low.
Selective reporting (reporting bias)	Unclear risk	Comment: the protocol is not publicly available for review, but the authors did provide the raw data, and there does not appear to be selective outcome reporting.

Papaioannou 2003a

Methods	Randomized controlled trial
Participants	<p># Randomized: 74 (37 intervention, 37 control)</p> <p>Lost to follow-up: 14 at 6 months, 17 in total at 12 months</p> <p>Age: 71.6 intervention, 72.2 control</p> <p>Country: Canada</p> <p>Inclusion criteria: women, at least one vertebral fracture (anterior height reduction of $\geq 15\%$), osteoporosis (lumbar spine T-score of ≤ -2.5)</p> <p>Exclusion criteria: resting heart rate ≥ 100 bpm, uncontrolled hypertension, secondary causes of bone loss, another diagnosis for back pain, unable to stand independently for 3 minutes</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> 3 times/week</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> 60 minutes</p> <p><i>Type:</i> stretching, strengthening with elastic band, aerobic activity</p> <p><i>Duration of intervention:</i> 52 weeks</p> <p><i>Supervisor:</i> Kinesiologist with MSc.</p> <p><i>Supervision/Setting:</i> intermittent home visits and phone calls, home exercise</p> <p><i>Progression of intensity:</i> yes</p> <p>Control: usual activities</p> <p>Duration of follow-up after baseline assessment: 52 weeks</p>
Outcomes	<p>Quality of life: Osteoporosis Quality of Life Questionnaire (disease-specific), Sickness Impact Profile</p> <p>Balance: postural sway</p> <p>Physical Performance: Timed Up and Go test</p> <p>Bone mineral density: DXA at lumbar spine and femoral neck</p>

Papaioannou 2003a (Continued)

*Webber 2003 report changes in vertebral height in a subgroup from this study - not included in the review

Notes	Adherence: 62% at 6 months, 46% at 12 months (those that were adherent to at least 3 times/week) Adverse events <i>Due to intervention:</i> - none reported <i>Unclear if due to intervention:</i> - fear of falling (n = 4) - pain or illness (n = 10) <i>Not due to intervention:</i> - 1 death Between-group comparison as actual difference or change scores: change scores	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: the randomization process was not described.
Allocation concealment (selection bias)	Unclear risk	Comment: the allocation concealment was not described.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Quote: "The research assistant who collected all data was blinded to intervention status." Comment: probably done.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: quality of life questionnaires were reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Comment: it is reported that an intention-to-treat analysis was performed, but it is not clear what was done with missing data.
Selective reporting (reporting bias)	Unclear risk	Comment: Webber 2003 published data on vertebral height changes from a subset of participants. The protocol is not publicly available for review.

Wang 2015

Methods	Randomized controlled trial
Participants	# Randomized: 138 (46 intervention, 46 control group 1, 46 control group 2)

Exercise for improving outcomes after osteoporotic vertebral fracture (Review)

Copyright © 2019 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Wang 2015 (Continued)

Lost to follow-up: 6 intervention, 10 control group 1, 8 control group 2

Age: 66.7 (65.8 intervention, 67.5 control group 1, 66.7 control group 2)

Country: China

Inclusion criteria: diagnosed with osteoporotic thoracic or lumbar vertebral compression fracture, presence of osteoporosis confirmed by DXA scan, apparent signs of pain, willing to undergo percutaneous kyphoplasty treatment, those without signs or symptoms of spinal cord or nerve root injury prior to surgery

Exclusion criteria: those with other accompanying serious bodily illnesses or lumbar disc herniations, malignant vertebral tumors or related spinal cord or nerve root injuries or symptoms, those with restrictions to percutaneous kyphoplasty surgery, those with mental illness or a history of mental illness in the family, those not approved by the hospital ethics committee or those unwilling to undergo the surgery

Interventions	Experimental <i>Frequency:</i> daily <i>Intensity:</i> not specified <i>Time:</i> 2 sets for 20 minutes <i>Type:</i> low back muscle training: three- and five-point support and swallow exercises, home exercise <i>Duration of intervention:</i> 52 weeks <i>Supervisor:</i> not specified <i>Supervision/Setting:</i> not specified <i>Progression of intensity:</i> not specified Control Group 1: conventional post-surgical treatment (Calcium and Calcitriol) plus Salmon Calcitonin Control Group 2: conventional post-surgical treatment (Calcium and Calcitriol) Duration of follow-up after baseline assessment: 52 weeks	
Outcomes	Pain: Visual Analogue Scale Self-reported physical function: Oswestry Disability Index Bone mineral density: DXA scan at lumbar spine	
Notes	Adherence: not specifically stated; the authors mentioned in the discussion that due to the difficulty of the low back training regimen, the patient compliance was relatively poor. Adverse events: none reported Between-group comparison as actual difference or change scores: change scores	
Risk of bias		
Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: participants were divided into three groups with a random number table with no further information on who conducted randomized and specific allocation ratio. There is not enough information provided in the methods to evaluate risk.

Wang 2015 (Continued)

Allocation concealment (selection bias)	Unclear risk	Comment: the allocation was conducted using a single-blind procedure with no further information on concealment. There is not enough information provided in the methods to evaluate risk.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Comment: there is not enough information provided in the methods to evaluate risk.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: pain and self-reported physical function were reported by participants who were aware of their group allocation.
Incomplete outcome data (attrition bias) All outcomes	High risk	Comment: initial recruited sample of 138 participants, yet the study only examined data of 114 patients that completed follow-up. There were no reasons provided for the lack of follow-up data in the excluded participants.
Selective reporting (reporting bias)	Unclear risk	Comment: there is not enough information provided in the methods to evaluate risk.

Yang 2007

Methods	Randomized controlled trial
Participants	<p># Randomized: 30 (15 intervention, 15 control)</p> <p>Lost to follow-up: none</p> <p>Age: 66.4 (67.4 intervention, 65.5 control)</p> <p>Country: China</p> <p>Inclusion criteria: women, more than one vertebral fracture (vertebrae compressed $\geq 20\%$), Chinese standard for primary osteoporosis diagnosis, willing to accept treatment</p> <p>Exclusion criteria: vertebral fractures due to trauma or malignancy, contraindications to exercise, cannot complete evaluation or treatment</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> twice per week</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> 10-30 minutes</p> <p><i>Type:</i> strengthening and flexibility exercises</p> <p><i>Duration of intervention:</i> 4 weeks</p> <p><i>Supervisor:</i> not reported</p> <p><i>Supervision/Setting:</i> unclear</p> <p><i>Progression of intensity:</i> no</p>

Yang 2007 (Continued)

Control: no description

Duration of follow-up after baseline assessment: 4 weeks

Outcomes	Pain: rating on a 100 mm visual analogue scale Physical performance: Time to sit up from a supine position, Timed Up and Go test
Notes	Adherence: 100% Adverse events: none reported Between-group comparison as actual difference or change scores: actual difference

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Comment: there is not enough information provided in the methods to evaluate risk.
Allocation concealment (selection bias)	Unclear risk	Comment: there is not enough information provided in the methods to evaluate risk.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: participants and personnel delivering the intervention were not blind to group allocation due to nature of intervention.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: outcome assessors were not blind to group allocation.
Blinding of subjective outcome assessment (detection bias)	High risk	Comment: pain was reported by participants who were aware of their group allocation. Outcome assessors were not blind to group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: all participants returned for follow-up.
Selective reporting (reporting bias)	Unclear risk	Comment: there is not enough information provided in the methods to evaluate risk. The protocol is not publicly available.

BMD: bone mineral density; bpm: beats per minute; DXA: dual energy X-ray absorptiometry; ITT: intention-to-treat; QUALEFFO-41: Quality of Life Questionnaire of the European Foundation for Osteoporosis

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Bada 2009	A brief descriptive review of treatments for vertebral fracture.
Barker 2014	Study protocol, ongoing; participants not all over the age of 40 years.
Bautmans 2010	Participants are a heterogeneous group of patients with and without vertebral fracture.

Study	Reason for exclusion
Bergland 2012	An abstract reporting data from Bergland 2011 and Olsen 2014 .
Borgo 2010	Participants who had a vertebral fracture at baseline were excluded from study participation.
Chan 2017	Vertebral fracture was not an inclusion criterion for the recruitment of participants. Absence of inactive control group.
Chen 2011	Control group participants were given non-compulsory activity and exercises (e.g. straight leg raising, Tai Chi and walking) post-operation.
Chen 2015	Absence of inactive control group.
Dusdal 2011	Systematic review, all studies that met our eligibility criteria were included in the original Cochrane Review (Giangregorio 2013).
Evstigneyeva 2013	An abstract with insufficient detail that included data from Evstigneeva 2016 . The authors provided the full-text article for the larger dataset that the abstract was based on and therefore, the abstract by Evstigneeva 2016 was included in our review.
Gandhi 2015	Analysis of barriers for outpatients recruitment in RCT studies.
Giangregorio 2014c	Study protocol for Giangregorio 2017 classified as awaiting assessment.
Hongo 2007	Participants are a heterogeneous group of patients with and without vertebral fracture. Although the outcomes and authors are similar to that reported in Miyakoshi 2010 , the interventions described in each trial are slightly different with respect to exercises chosen, follow-up time, and exercise frequency so it is likely that they are independent trials. Three attempts to contact the authors via email were made with no response.
Jensen 2012	No information about vertebral fracture as an inclusion criterion.
Kaijser Alin 2016	Vertebral fracture was not an inclusion criterion for the recruitment of participants.
Karakasidou 2013	Control group participants were given non-compulsory activity exercises.
Liu 2017	Not an RCT study, but a systematic review regarding the effects of exercise on people with vertebral fracture.
Lord 1996	Participants did not have a vertebral fracture.
Malmros 1999	Same study as Malmros 1998 but in Danish.
Marcu 2015	Vertebral fracture was not an inclusion criterion for the recruitment of participants.
Mihailov 2011	An abstract with insufficient details and no estimates of effect were reported in the article. We contacted the authors to get access to all unpublished outcomes from the dataset, but were not able to obtain them in time for this version of the review.
Miyakoshi 2010	No information about control group activities to confirm that they did not receive an intervention. No information about randomization to confirm it was a randomized trial. No estimates of effect provided (P values only).
Nemes 2014	An abstract with insufficient details. We contacted the authors to get access to all unpublished outcomes from the dataset, but were not able to obtain them in time for this version of the review.

Study	Reason for exclusion
Rapp 2011	Descriptive summary of another RCT with insufficient information on trial and no data reported in the article.
Rittchen 1991	Not a randomized controlled trial, but a descriptive paper about the management of patients with acute fractures.
Schwinning 1992	Not a randomized controlled trial, but a descriptive paper about the management of patients with acute fractures.
Shipp 2004	An abstract reporting data from Gold 2004 . There is no formal between-group comparison in this abstract, but there is one sentence implying that a between-group difference in six-minute walk distance was not significant, which we have included in the review. The focus of the abstract was to investigate whether changes in six-minute walk distance was related to changes in other measures of disability. We were in contact with the authors of Gold 2004 to get access to all unpublished outcomes from the larger dataset, but were not able to obtain them in time for this version of the review.
Shipp 2007	An abstract reporting previously published data from Gold 2004 , as well as unpublished data on one outcome. We were in contact with the authors of Gold 2004 to get access to all unpublished outcomes from the larger dataset, but were not able to obtain them in time for this version of the review. We did report the findings for the unpublished outcome in our review, but there was not enough information in the abstract to warrant its inclusion independently.
Sinaki 1984	Not a randomized controlled trial.
Sinaki 1995	Not a randomized controlled trial.
Smith 1998	Not all participants had a vertebral fracture. Participants had several types of fracture e.g. vertebral, hip, radius, clavicle. No information on randomization process provided in published report, or in methodology section of unpublished report that we obtained from one of the authors. Therefore it is not possible to confirm that it was truly a randomized trial.
Stadhouder 2009	Includes patients with acute traumatic vertebral fracture only. Includes patients who are less than 40 years of age. Comparison group is using a brace.
Svensson 2017	Not a RCT study but a review on supportive intervention for women with vertebral fracture.
Webber 2003	A secondary analysis of data from a subgroup of patients in Papaioannou 2003a , using an outcome that was not specified in the protocol, namely vertebral height as a continuous measure.
Zhong 2012	Control group participants were given non-compulsory activity and exercises (e.g. leg exercises, walking) post-operation.

RCT: randomized controlled trial

Characteristics of studies awaiting assessment *[ordered by study ID]*

[Bergstrom 2016](#)

Methods	Randomized controlled trial
Participants	36 women with osteoporosis and vertebral fracture Country: Sweden

Bergstrom 2016 (Continued)

Interventions	<p>Experimental</p> <p><i>Frequency:</i> 2 times per week</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> not specified</p> <p><i>Type:</i> back muscle training program</p> <p><i>Duration of intervention:</i> 16 weeks</p> <p><i>Supervisor:</i> physical therapist</p> <p><i>Supervision/Setting:</i> continual supervision in a group exercise class</p> <p><i>Progression of intensity:</i> unclear</p> <p>Control: not described</p> <p>Duration of follow-up after baseline assessment: 16 weeks</p>
Outcomes	Back extensor muscle strength, kyphosis, height, and thoracic expansion
Notes	Abstract presented at World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases 2016; Original trial included in review (Bergstrom 2011), will assess for extraction of other outcomes.

Giangregorio 2017

Methods	Pilot feasibility study with 1:1 randomization
Participants	<p>Women aged 65 years or older with vertebral fracture at 5 Canadian and 2 Australian centers will be recruited</p> <p>Countries: Canada and Australia</p> <p>Inclusion criteria: women, aged 65 years or older, have radiographic evidence of a nontraumatic fracture of at least 1 vertebrae between T4 and L4 (vertebrae compressed $\geq 25\%$)</p> <p>Exclusion criteria: index vertebral fracture due to trauma, not able to communicate in English, on dialysis, receiving palliative care, current or prior cancer (except basal cell carcinoma), clinically significant kidney, liver, or intestinal disease, exercise participation ≥ 3 times per week that addressed ≥ 2 of 5 domains in the Build Better Bones With Exercise exercise prescription, progressive neurological disorder or progressive disorder likely to prevent study completion, unable to stand or walk with or without a gait aid, impaired capacity to give informed consent (positive Mini-Cog score and cannot recall what they are being asked to do during the consent process), or contraindication to exercise as determined by a physician.</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> at least 3 times/week</p> <p><i>Intensity:</i> lower- and upper-extremity strengthening: use body weight, the floor or wall, or a light-weight TheraBand exercise band as resistance- maximum resistance with which 10 repetitions can be completed with good form (10-RM) and begin with 2 sets of 8 to 10 repetitions at 10-RM; posture training: hold each position for 3 seconds and repeat 3 times; balance training: difficulty will be individually tailored; aerobic exercise: 5 to 8 on the Borg Rating of Perceived Exertion Scale</p> <p><i>Time:</i> at least one exercise from each domain; aerobic training: begin with 10-30 minutes; duration is progressively increased</p>

Giangregorio 2017 (Continued)

Type: Build Better Bones with Exercise exercise prescription, including exercises focused on lower and upper extremity strengthening, balance, posture, and endurance training, and the Motivation to Move Program (a behavioral counseling guide informed by motivating interviewing and the health action process approach)

Duration of intervention: 52 weeks

Supervisor: physical therapist

Supervision/Setting: biweekly phone calls and 4 home visits (first 2 months), and monthly phone calls and 2 home visits (after 2 months), home exercise

Progression of intensity: yes

Control: same number of home visits and phone calls as intervention group (equal attention) with physical therapist; focused on health and social discussion

Duration of follow-up after baseline assessment: 52 weeks

Outcomes	<p>Primary outcomes: recruitment, retention, and adherence</p> <p>Secondary outcomes: physical performance (lower extremity strength and balance), posture, falls, quality of life (disease-specific, generic), pain, fall self-efficacy, behavior change variables, intervention cost, fractures, and adverse events</p>
Notes	<p>Abstract presented at the American Society for Bone and Mineral Research Conference in 2017; ClinicalTrials.gov identifier: NCT01761084</p>

Giangregorio 2018

Methods	Pilot feasibility study with 1:1 randomization
Participants	<p>Women aged 65 years or older with vertebral fracture at 5 Canadian and 2 Australian centers will be recruited</p> <p>Countries: Canada and Australia</p> <p>Inclusion criteria: women, aged 65 years or older, have radiographic evidence of a nontraumatic fracture of at least 1 vertebrae between T4 and L4 (vertebrae compressed $\geq 25\%$)</p> <p>Exclusion criteria: index vertebral fracture due to trauma, not able to communicate in English, on dialysis, receiving palliative care, current or prior cancer (except basal cell carcinoma), clinically significant kidney, liver, or intestinal disease, exercise participation ≥ 3 times per week that addressed ≥ 2 of 5 domains in the Build Better Bones With Exercise exercise prescription, progressive neurological disorder or progressive disorder likely to prevent study completion, unable to stand or walk with or without a gait aid, impaired capacity to give informed consent (positive Mini-Cog score and cannot recall what they are being asked to do during the consent process), or contraindication to exercise as determined by a physician.</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> at least 3 times/week</p> <p><i>Intensity:</i> lower- and upper-extremity strengthening: use body weight, the floor or wall, or a light-weight TheraBand exercise band as resistance- maximum resistance with which 10 repetitions can be completed with good form (10-RM) and begin with 2 sets of 8 to 10 repetitions at 10-RM; posture training: hold each position for 3 seconds and repeat 3 times; balance training: difficulty will be individually tailored; aerobic exercise: 5 to 8 on the Borg Rating of Perceived Exertion Scale</p>

Giangregorio 2018 (Continued)

Time: at least one exercise from each domain; aerobic training: begin with 10-30 minutes; duration is progressively increased

Type: Build Better Bones with Exercise exercise prescription, including exercises focused on lower and upper extremity strengthening, balance, posture, and endurance training, and the Motivation to Move Program (a behavioral counseling guide informed by motivating interviewing and the health action process approach)

Duration of intervention: 52 weeks

Supervisor: physical therapist

Supervision/Setting: biweekly phone calls and 4 home visits (first 2 months), and monthly phone calls and 2 home visits (after 2 months), home exercise

Progression of intensity: yes

Control: same number of home visits and phone calls as intervention group (equal attention) with physical therapist; focused on health and social discussion

Duration of follow-up after baseline assessment: 52 weeks

Outcomes	<p>Primary outcomes: recruitment, retention, and adherence</p> <p>Secondary outcomes: physical performance (lower extremity strength and balance), posture, falls, quality of life (disease-specific, generic), pain, fall self-efficacy, behavior change variables, intervention cost, fractures, and adverse events</p>
Notes	ClinicalTrials.gov identifier: NCT01761084

Pompa 2016

Methods	Randomized controlled trial
Participants	<p>Women with dorsal hyperkyphosis due to one or more vertebral wedging by postmenopausal osteoporosis (T-score<-2.5)</p> <p>Country: Italy</p> <p>Inclusion criteria: female; aged between 55 and 65 years, postmenopausal osteoporosis; one or more vertebral fractures.</p> <p>Exclusion criteria: recent fractures (MRI negative for bone edema); usual program of exercise/rehabilitation; neurological or orthopedic-associated serious diseases</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> 1 session every 5 days</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> 1 hour per session</p> <p><i>Type:</i> postural and respiratory rehabilitation focused on the mobilization in elongation of the paravertebral and respiratory muscles, in addition to other exercises that patients were educated to daily</p> <p><i>Duration of intervention:</i> 8 weeks</p> <p><i>Supervisor:</i> physical therapist</p>

Pompa 2016 (Continued)

Supervision/Setting: not specified

Progression of intensity: not specified

Control: did not perform any kind of rehabilitation

Duration of follow-up after baseline assessment: 8 weeks

Outcomes	Numerical rating scale for pain; distance hands to floor (in flexion and side bending) for evaluation of spinal mobility; distance tragus to wall and plumb line to C7 method for evaluation of dorsal kyphosis; QUALEFFO (disease-specific quality of life)
Notes	Abstract presented at World Congress on Osteoporosis, Osteoarthritis and Musculoskeletal Diseases 2016

MRI: magnetic resonance imaging; QUALEFFO-41: Quality of Life Questionnaire of the European Foundation for Osteoporosis

Characteristics of ongoing studies [ordered by study ID]

Stanghelle 2018

Trial name or title	Effect of a resistance and balance exercise programme for women with osteoporosis and vertebral fracture: study protocol for a randomized controlled trial
Methods	Parallel-group single-blinded 1:1 randomized controlled trial
Participants	<p>Women aged 65 years or older with vertebral fracture</p> <p>Country: Norway</p> <p>Inclusion criteria: women over 65 years, living at home, able to walk with or without walking aids; having osteoporosis verified by DXA scan (-2.5); at least one vertebral fracture, verified by X-ray</p> <p>Exclusion criteria: health conditions that make it unsafe to exercise; not able to understand and speak Norwegian adequately to participate in the intervention</p>
Interventions	<p>Experimental</p> <p><i>Frequency:</i> at least 2 times/week</p> <p><i>Intensity:</i> not specified</p> <p><i>Time:</i> not specified</p> <p><i>Type:</i> group circuit exercise program, focusing on lower limb strength, back strength and static and dynamic balance</p> <p><i>Duration of intervention:</i> 12 weeks</p> <p><i>Supervisor:</i> not specified</p> <p><i>Supervision/Setting:</i> center-based, group exercise</p> <p><i>Progression of intensity:</i> unclear</p> <p>Control: instructed to "live as usual"</p> <p>Duration of follow-up after baseline assessment: 24 weeks</p>
Outcomes	Primary outcome: 10-metre habitual walking speed

Stanghelle 2018 (Continued)

Secondary outcomes: quality of life, fitness, balance, grip strength, falls self-efficacy, physical activity, physical performance

Starting date	January 2016
Contact information	Brita Stanghelle, PhD-student, Oslo and Akershus University college of Applied Sciences, Oslo, Norway, 0130 +47 97514751 brita.stanghelle@hioa.no
Notes	Completed recruitment; ClinicalTrials.gov identifier: NCT02781974

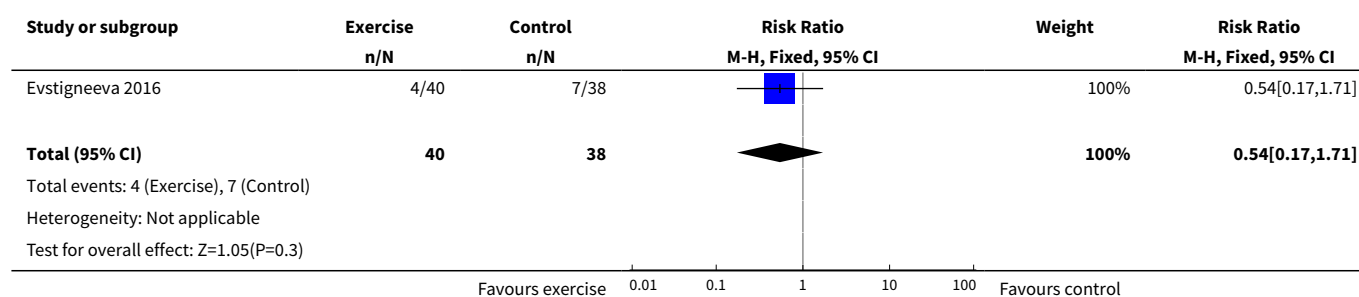
DXA: dual energy X-ray absorptiometry

DATA AND ANALYSES

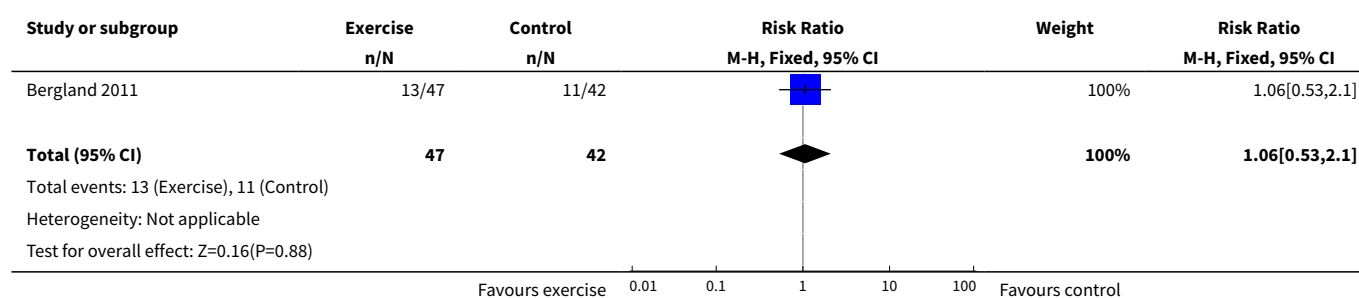
Comparison 1. Exercise versus Control (After 52 weeks)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Incident Fragility Fracture	1	78	Risk Ratio (M-H, Fixed, 95% CI)	0.54 [0.17, 1.71]
2 Incident Falls	1	89	Risk Ratio (M-H, Fixed, 95% CI)	1.06 [0.53, 2.10]

Analysis 1.1. Comparison 1 Exercise versus Control (After 52 weeks), Outcome 1 Incident Fragility Fracture.



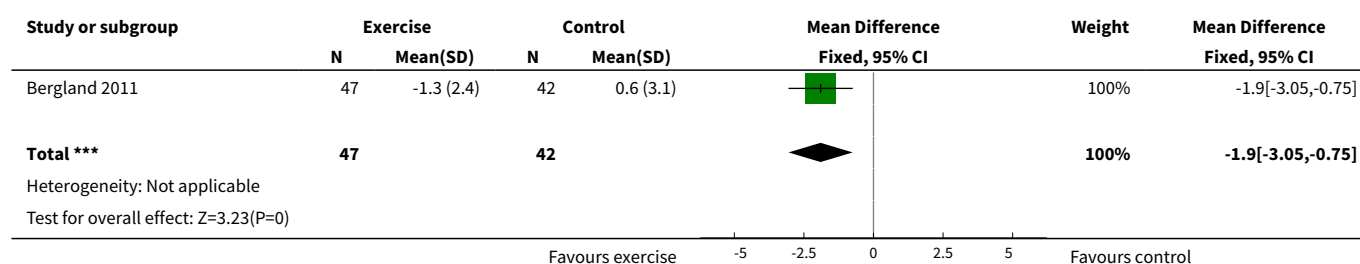
Analysis 1.2. Comparison 1 Exercise versus Control (After 52 weeks), Outcome 2 Incident Falls.



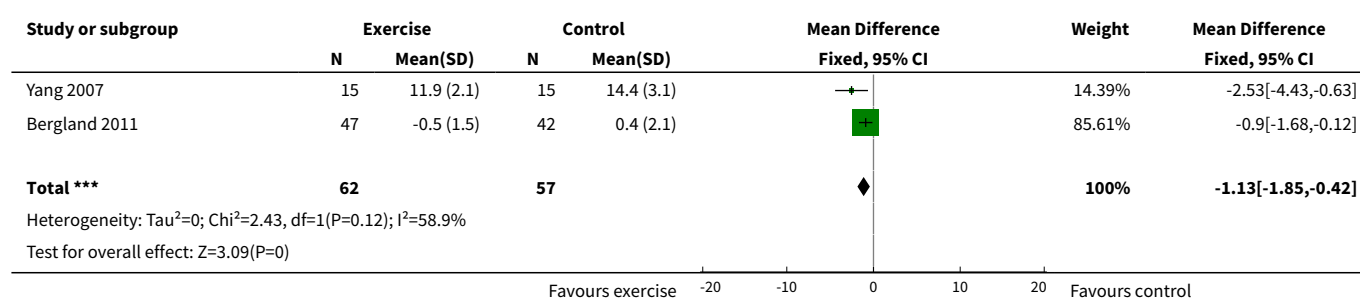
Comparison 2. Exercise versus Control (After 12 weeks)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Maximal Walking Speed over 20 meters	1	89	Mean Difference (IV, Fixed, 95% CI)	-1.9 [-3.05, -0.75]
2 Timed Up and Go	2	119	Mean Difference (IV, Fixed, 95% CI)	-1.13 [-1.85, -0.42]
3 QUALEFFO-41 Physical Function Score	2	109	Mean Difference (IV, Fixed, 95% CI)	-2.84 [-5.57, -0.11]
4 QUALEFFO-41 Total Score	2	109	Mean Difference (IV, Fixed, 95% CI)	-3.24 [-6.05, -0.43]

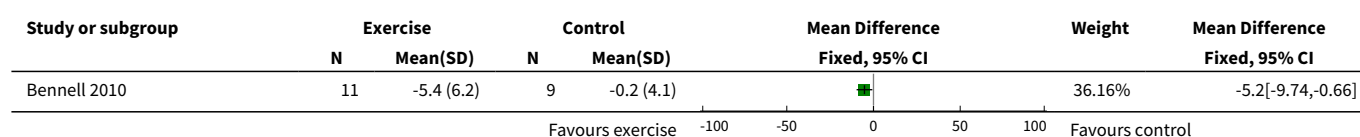
Analysis 2.1. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 1 Maximal Walking Speed over 20 meters.

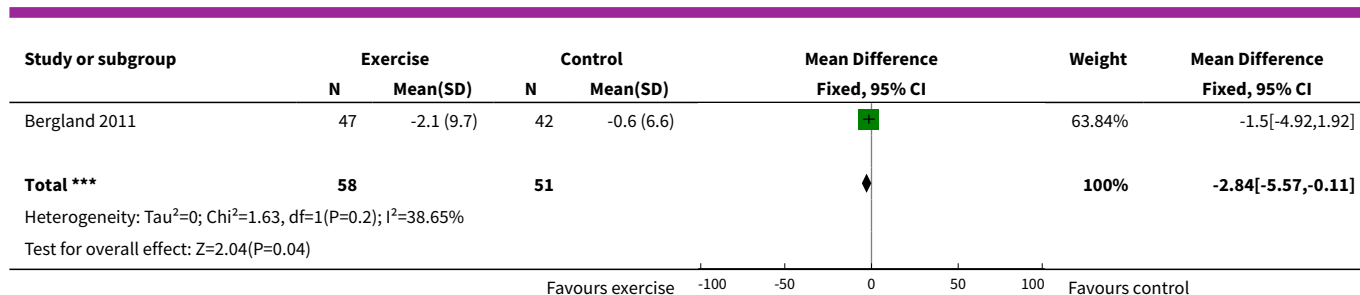


Analysis 2.2. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 2 Timed Up and Go.

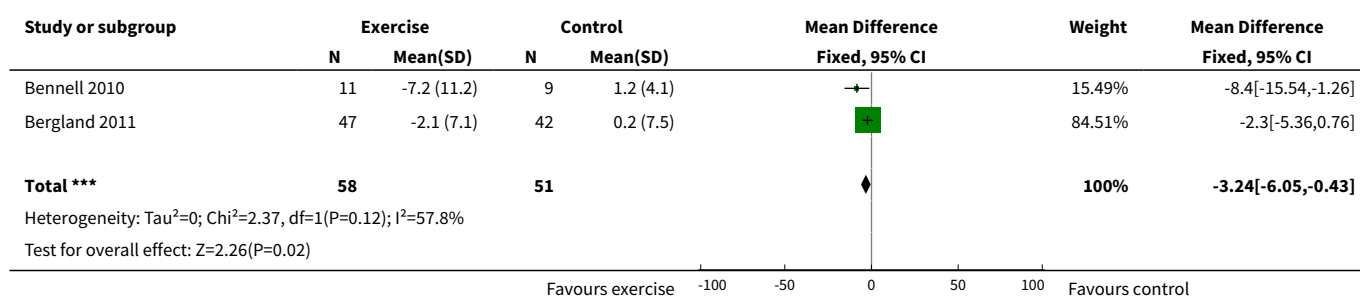


Analysis 2.3. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 3 QUALEFFO-41 Physical Function Score.





Analysis 2.4. Comparison 2 Exercise versus Control (After 12 weeks), Outcome 4 QUALEFFO-41 Total Score.



ADDITIONAL TABLES

Table 1. Methodological quality assessment scheme (adapted from Cochrane's tool for assessing risk of bias)

Domain	Score	Domain Description	Comments
Was the allocation sequence adequately generated?	YES	There is a random component in sequence generation	
	UNCLEAR	Method of randomization not stated or unclear	
	NO	Quasi-randomized, nonrandom component in sequence generation	
Was allocation adequately concealed prior to or during randomization?	YES	Participants/investigators could not foresee assignments	
	UNCLEAR	Method of allocation concealment not stated or unclear	
	NO	Participants/investigators could possibly foresee assignments, quasi-randomized	
Were outcome assessors blinded to treatment status?	YES	Blinding of outcome assessment, or outcomes unlikely to be affected by lack of blinding	
	UNCLEAR	Insufficient information to determine if blinding did or did not occur	
	NO	No blinding, incomplete blinding, chance blinding could be broken, AND lack of blinding is likely to introduce bias	
Were incomplete outcome data adequately addressed?	YES	No missing data, or missing data are: balanced across groups, unlikely to affect outcome, imputed, ITT analysis	
	UNCLEAR	Insufficient information about attrition/exclusions	

Table 1. Methodological quality assessment scheme (adapted from Cochrane's tool for assessing risk of bias) (Continued)

	NO	Missing data likely to affect outcome or be related to outcome, as-treated analysis, inappropriate imputation
Are reports of the study free of selective outcome reporting?	YES	Protocol is available and measurement methods for pre-specified outcomes defined and reported as defined, or key expected outcomes have been defined and reported
	UNCLEAR	
	NO	Insufficient information to judge whether or not selective outcome reporting has occurred Incomplete or absent reporting, or key outcomes not reported that would be expected, measurement methods not specified

ITT: intention-to-treat

Table 2. Adverse events reported in exercise trials in individuals with vertebral fracture

Adverse Event	Number of Incidences Per Study	Number of Incidences Per Group	Due to Intervention	Cause	Resulted in Study Withdrawal	Study
Death	1	Unknown	No	Unknown	Yes	Papaioannou 2003a
Fracture of costal cartilage	1	Exercise Group: 1	Yes	Prone Exercise	Unknown	Gold 2004
Rib Fracture	1	Exercise Group: 1	Yes	Rolling from supine to prone	Unknown	Gold 2004
Vertebral Fracture	4	Exercise Group: 2 Control Group: 2	No	Unknown	No	Evstigneeva 2016
Hip Fracture	1	Unknown	No	Study physical examination	Unknown	Gold 2004
Metatarsal Fracture	1	Unknown	No	Study assessment; 2lb weight fell on foot	Unknown	Gold 2004
Non-vertebral Fracture	7	Exercise Group: 2 Control Group: 5	No	Unknown	No	Evstigneeva 2016
Myeloma diagnosis	1	Exercise Group: 1	No		No	Bergstrom 2011
Knee Pain	1	Exercise Group: 1	Yes	Exercise in knee-wrist position	No	Evstigneeva 2016
Pain	4	Exercise Group: 2 (Gold 2004), 1 (Bergstrom 2011) Control Group: 1 (Bergstrom 2011)	Unknown	Soft tissue origin	Unclear - resulted in missed classes	Gold 2004 ; Bergstrom 2011

Table 2. Adverse events reported in exercise trials in individuals with vertebral fracture (Continued)

Pain or illness	10	Unknown	Unknown	Unknown	Yes	Papaioannou 2003a
Pain or injury	6	Exercise Group: 5 Control Group: 1	Unknown	Unknown	No	Bennell 2010
Irritation to tape	1	Exercise Group: 1	Yes	Reaction to tape material	No	Bennell 2010
Fear of falling or fall	4	Unknown	Unknown	Unknown	Yes	Papaioannou 2003a
Undescribed adverse events that caused study withdrawal	5	Unknown	Author indicated they were unrelated	Unknown	Yes	Malmros 1999

The adverse events here are reported in the results of each study, but not all studies mentioned adverse events. There was no clear indication in any of the studies that adverse events were systematically monitored.

Table 3. Sensitivity analysis - Exercise versus Control (After 4 to 12 weeks)

Outcome or subgroup	Studies	Participants	Statistical Method	Effect Estimate
Timed Up and Go - Without Bennell 2010	2	119	Mean Difference (IV, Fixed, 95% CI)	- 1.13 [-1.85, -0.42]
Timed Up and Go - With Bennell 2010	3	139	Mean Difference (IV, Fixed, 95% CI)	- 1.09 [-1.78, -0.40]

Sensitivity analysis: [Bergland 2011](#) and [Yang 2007](#) pooled without and with [Bennell 2010](#) (excluded from the first edition of this review because there was no comparison group that received the same intervention with the exception of exercise) to further assess the effects of exercise on Timed Up and Go test performance.

APPENDICES

Appendix 1. MEDLINE search strategy

1. Spinal Fractures/
2. exp Spine/
3. (spine\$ or spinal or vertebra\$).tw.
4. exp Fractures, Bone/
5. fractur\$.tw.
6. 2 or 3
7. 4 or 5
8. 6 and 7
9. Vertebroplasty/
10. vertebroplast\$.tw.

11. kyphoplast\$.tw.
12. or/1,8-11
13. exp EXERCISE/
14. Physical Exertion/
15. exp Sports/
16. exp PLIABILITY/
17. exertion\$.tw.
18. exercis\$.tw.
19. sport\$.tw.
20. exp Physical Therapy Modalities/
21. ((physical or motion) adj5 (fitness or therap\$)).tw.
22. (physical\$ adj2 endur\$).tw.
23. manipulat\$.tw.
24. Muscle Strength/
25. ((muscle or strength\$ or extensor or weight) adj (training or resistance)).tw.
26. resistance training.tw.
27. exp rehabilitation/
28. rh.fs.
29. rehab\$.tw.
30. kinesiotherap\$.tw.
31. or/13-30
32. 12 and 31
33. randomized controlled trial.pt.
34. controlled clinical trial.pt.
35. randomized.ab.
36. placebo.ab.
37. drug therapy.fs.
38. randomly.ab.
39. trial.ab.
40. groups.ab.
41. or/33-40
42. (animals not (humans and animals)).sh.
43. 41 not 42
44. 32 and 43

Appendix 2. Embase search strategy

1. exp spine fracture/
2. exp spine/
3. (spine\$ or spinal or vertebra\$).tw.
4. exp Fracture/
5. fractur\$.tw.
6. 2 or 3
7. 4 or 5
8. 6 and 7
9. kyphoplasty/ or percutaneous vertebroplasty/
10. vertebroplast\$.tw.
11. kyphoplast\$.tw.
12. or/1,8-11
13. exp EXERCISE/
14. pliability/
15. exertion\$.tw.
16. exercis\$.tw.
17. exp physiotherapy/
18. ((physical or motion) adj5 (fitness or therap\$)).tw.
19. (physical\$ adj2 endur\$).tw.
20. manipulat\$.tw.
21. muscle strength/
22. ((muscle or strength\$ or extensor or weight) adj (training or resistance)).tw.
23. resistance training.tw.
24. exp Rehabilitation/
25. rh.fs.
26. rehab\$.tw.
27. kinesiotherap\$.tw.
28. or/13-27
29. 12 and 28
30. random\$.ti,ab.
31. factorial\$.ti,ab.
32. (crossover\$ or cross over\$ or cross-over\$).ti,ab.
33. placebo\$.ti,ab.
34. (doubl\$ adj blind\$).ti,ab.

35. (singl\$ adj blind\$).ti,ab.
36. assign\$.ti,ab.
37. allocat\$.ti,ab.
38. volunteer\$.ti,ab.
39. crossover procedure.sh.
40. double blind procedure.sh.
41. randomized controlled trial.sh.
42. single blind procedure.sh.
43. or/30-42
44. exp animal/ or nonhuman/ or exp animal experiment/
45. exp human/
46. 44 and 45
47. 44 not 46
48. 43 not 47
49. 29 and 48

Appendix 3. *The Cochrane Library* search strategy

- #1 MeSH descriptor Spinal Fractures, this term only
- #2 MeSH descriptor Spine explode all trees
- #3 (spine* or spinal or vertebra*):ti,ab
- #4 MeSH descriptor Fractures, Bone explode all trees
- #5 fractur*:ti,ab
- #6 (#2 OR #3)
- #7 (#4 OR #5)
- #8 (#6 AND #7)
- #9 MeSH descriptor Vertebroplasty explode all trees
- #10 vertebroplast*:ti,ab
- #11 kyphoplast*:ti,ab
- #12 (#1 OR #8 OR #9 OR #10 OR #11)
- #13 MeSH descriptor Exercise explode all trees
- #14 MeSH descriptor Physical Exertion, this term only
- #15 MeSH descriptor Sports explode all trees
- #16 MeSH descriptor Pliability, this term only
- #17 exertion*:ti,ab
- #18 exercis*:ti,ab
- #19 sport*:ti,ab

- #20 MeSH descriptor Physical Therapy Modalities explode all trees
- #21 ((physical or motion) near/5 (fitness or therap*)):ti,ab
- #22 (physical* near/2 endur*):ti,ab
- #23 manipul*:ti,ab
- #24 MeSH descriptor Muscle Strength, this term only
- #25 ((muscle or strength* or extensor or weight) next (training or resistance)):ti,ab
- #26 resistance next training:ti,ab
- #27 MeSH descriptor Rehabilitation explode all trees
- #28 Any MeSH descriptor with qualifier: RH
- #29 rehab*:ti,ab
- #30 kinesiotherap*:ti,ab
- #31 (#13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29 OR #30)
- #32 (#12 AND #31)

Appendix 4. CINAHL search strategy

S55 S41 and S54

S54 S42 or S43 or S44 or S45 or S46 or S47 or S48 or S49 or S50 or S51 or S52 or S53

S53 TI Allocat* random* or AB Allocat* random*

S52 (MH "Quantitative Studies")

S51 (MH "Placebos")

S50 TI Placebo* or AB Placebo*

S49 TI Random* allocat* or AB Random* allocat*

S48 (MH "Random Assignment")

S47 TI Randomi?ed control* trial* or AB Randomi?ed control* trial*

S46 AB singl* blind* or AB singl* mask* or AB doub* blind* or AB doubl* mask* or AB trebl* blind* or AB trebl* mask* or AB tripl* blind* or AB tripl* mask*

S45 TI singl* blind* or TI singl* mask* or TI doub* blind* or TI doubl* mask* or TI trebl* blind* or TI trebl* mask* or TI tripl* blind* or TI tripl* mask*

S44 TI clinical* trial* or AB clinical* trial*

S43 PT clinical trial

S42 (MH "Clinical Trials+")

S41 S12 and S40

S40 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 or S36 or S37 or S38 or S39

S39 ti kinesiotherap* or ab kinesiotherap*

S38 ti rehab* or ab rehab*

S37 (MH "Rehabilitation+")

S36 ti resistance training or ab resistance training

S35 ti weight resistance or ab weight resistance

S34 ti weight training or ab weight training
S33 ti extensor resistance or ab extensor resistance

S32 ti extensor training or ab extensor training

S31 ti strength resistance or ab strength resistance

S30 ti strength training or ab strength training

S29 ti muscle resistance or ab muscle resistance

S28 ti muscle training or ab muscle training

S27 (MH "Muscle Strength")

S26 ti manipul* or ab manipul*

S25 ti physical* N2 endur* or ab physical* N2 endur*

S24 ti motion N5 therap* or ab motion N5 therap*

S23 ti motion N5 fitness or ab motion N5 fitness

S22 ti physical N5 therap* or ab physical N5 therap*

S21 ti physical N5 fitness or ab physical N5 fitness

S20 (MH "Physical Therapy+")

S19 ti sport* or ab sport*

S18 ti exercis* or ab exercis*

S17 ti exertion* or ab exertion*

S16 (MH "Pliability")

S15 (MH "Sports+")

S14 (MH "Exertion+")

S13 (MH "Exercise+")

S12 S1 or S9 or S10 or S11

S11 ti kyphoplast* or ab kyphoplast*

S10 ti vertebroplast* or ab vertebroplast*

S9 S7 and S8

S8 S5 or S6

S7 S2 or S3 or S4

S6 ti fractur* or ab fractur*

S5 (MH "Fractures+")

S4 ab spine* or spinal or vertebra*

S3 ti spine* or spinal or vertebra*

S2 (MH "Spine+")

S1 (MH "Spinal Fractures+")

Appendix 5. AMED search strategy

1. spinal fractures/

2. exp spine/

3. (spine\$ or spinal or vertebra\$).tw.

4. exp fractures bone/

5. fractur\$.tw.

6. 2 or 3

7. 4 or 5

8. 6 and 7

9. vertebroplast\$.tw.
10. kyphoplast\$.tw.
11. or/1,8-10
12. exp exercise/
13. exp exertion/
14. exp sports/
15. pliability/
16. exertion\$.tw.
17. exercis\$.tw.
18. sport\$.tw.
19. Physiotherapy/
20. exp physical therapy modalities/
21. ((physical or motion) adj5 (fitness or therap\$)).tw.
22. (physical\$ adj2 endur\$).tw.
23. manipulatt\$.tw.
24. muscle strength/
25. ((muscle or strength\$ or extensor or weight) adj (training or resistance)).tw.
26. resistance training.tw.
27. rehabilitation/
28. rehab\$.tw.
29. kinesiotherap\$.tw.
30. or/12-29
31. 11 and 30

Appendix 6. PEDro search strategy

Search 1

Spinal fracture* in Abstract & Title AND

Fitness training in Therapy AND

Clinical trial in Method

Search 2

Spinal fracture* in Abstract & Title AND

Strength training in Therapy AND

Clinical trial in Method

Search 3

Vertebroplast* in Abstract & Title AND

Fitness training in Therapy AND

Clinical trial in Method

Search 4

Vertebroplast* in Abstract & Title AND

Strength training in Therapy AND

Clinical trial in Method

Search 5

kyphoplast* in Abstract & Title AND

Fitness training in Therapy AND

Clinical trial in Method

Search 6

kyphoplast* in Abstract & Title AND

Strength training in Therapy AND

Clinical trial in Method

Appendix 7. World Health Organization International Clinical Trials Registry Platform search strategy

Search 1

Spinal fracture* in Title

Search 2

Vertebroplast* in Title

Search 3

kyphoplast* in Title

Appendix 8. ClinicalTrials.gov search strategy

Completed | Interventional Studies | vertebral fracture

WHAT'S NEW

Date	Event	Description
20 March 2019	New search has been performed	Search updated to March 2019. The number of new references after de-duplication was 461. Thirty three new records were assessed for eligibility and 31 records were excluded.
20 March 2019	New citation required and conclusions have changed	<p>This review represents an update of a Cochrane Review first published in 2013 (Giangregorio 2013).</p> <ul style="list-style-type: none"> The list of outcomes was revised to adhere to the reporting recommendations of Cochrane Musculoskeletal (Ghogomu 2014) and to match our major outcomes with the 'Summary of findings' table. Fear of falling and patient global assessment of success were added to the list of outcome measures. Two new trials were eligible for inclusion in the review (Evstigneeva 2016 and Wang 2015; 216 participants). We also extracted data on the effect of exercise on self-reported falls and fear of falling from a secondary analysis by Olsen 2014,

Date	Event	Description
		<p>which were not reported in the original trial by Bergland 2011 (included in the first edition of this review).</p> <ul style="list-style-type: none"> Conclusions changed to suggest: there was moderate-quality evidence that exercise can improve physical performance, specifically Timed Up and Go test, in individuals with a history of vertebral fracture (downgraded due to limitations of studies).

HISTORY

Protocol first published: Issue 7, 2010

Review first published: Issue 1, 2013

Date	Event	Description
16 June 2010	Amended	CMSS ID: C193-P

CONTRIBUTIONS OF AUTHORS

LG, AP, and LT conceived the idea for the review. LG, NM, AP, and LT contributed to the design of the protocol, in addition to Ann Cranney and Richard Cook, who were not involved in the preparation of the full report. LG, JG, NM, MP, and JT reviewed full texts for eligibility and independently extracted data from included trials. LG, JG, and MP were responsible for writing the review and incorporating all authors' feedback.

DECLARATIONS OF INTEREST

J. Gibbs: None.

N. MacIntyre: None.

M. Ponzano: None.

J. Templeton: None.

L. Thabane: is a member of an Advisory Board for Network Meta-analysis for GSK.

A. Papaioannou: is or has been a consultant, or on a speaker's bureau, or received unrestricted grants for the following: Amgen, Eli Lilly, Merck Frosst Canada, Novartis, Pfizer, Warner Chilcott.

L. Giangregorio: has consulted for ICON on behalf of Eli Lilly related to physical performance and osteoporosis.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- Canadian Institutes of Health Research, Canada.

Lora Giangregorio was the recipient of the Randomized Controlled Trials Mentoring Program Award (2008-2010) and is a CIHR New Investigator (2011-2016).

Norma MacIntyre was the recipient of the Randomized Controlled Trials Mentoring Program Award (2010-2012).

Jenna Gibbs is the recipient of a Canadian Institutes of Health Research Fellowship.

- Canadian Institutes of Health Research/Eli Lilly, Canada.

A. Papaioannou holds the Canadian Institutes for Health Research/Eli Lilly Chair in Osteoporosis, Falls and Fracture Prevention.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We reorganized the list of outcomes to adhere to the reporting recommendations of Cochrane Musculoskeletal ([Ghogomu 2014](#)) and match our major outcomes with the 'Summary of findings' table.

We initially removed fear of falling from the list because none of the studies assessed fear of falling in the first published version of this Cochrane Review. However, one study was identified in the new literature search that reported on the effect of exercise on fear of falling and therefore, we included fear of falling as an outcome in this updated version of the Cochrane Review. Patient global assessment of success was added as a minor outcome to adhere to the reporting recommendations of Cochrane Musculoskeletal ([Ghogomu 2014](#)).

In most cases we chose not to perform any meta-analyses because of the limited number of studies, the availability of data for pooling or the variability across studies with respect to outcomes chosen. A sensitivity analysis was conducted to compare the Timed Up and Go test results with and without [Bennell 2010](#) (excluded in the first published version of this Cochrane Review because there was no comparison group that received the same intervention with the exception of exercise).

INDEX TERMS

Medical Subject Headings (MeSH)

*Exercise Therapy [methods]; Exercise [physiology]; Osteoporotic Fractures [*therapy]; Postural Balance; Quality of Life; Randomized Controlled Trials as Topic; Spinal Fractures [*therapy]; Time and Motion Studies

MeSH check words

Humans